THE APPLICATION OF THE STIMULUS EQUIVALENCE PARADIGM TO COGNITIVELY COMPLEX ACADEMIC TASKS AND CHILDREN WITH LIMITED ENGLISH REPERTOIRES

BY

JOLENEA B. FERRO

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1990

ACKNOWLEDGEMENTS

There are many people whose support, encouragement, and love have sustained me for the past three years. First and foremost are my children, who were patient when I was not, understanding when my work kept me glued to the computer instead of attending to them, and helpful in ways I would not have thought possible.

I am also indebted to the Special Education faculty, whose faith in me never waivered. I am particularly grateful to my doctoral committee. Dr. William Wolking, my committee chair, provided support, encouragement, and excellent problem-solving skills when I thought there were no solutions. Drs. William Reid, Cary Reichard, and Roy Bolduc were always available to help and encourage me. Dr. Brian Iwata helped me look at my research critically and pushed me to become a better researcher in the process. I could not have asked for a more expert or sensitive committee.

Finally, I extend my love and gratitude to my parents, who believed I could do anything, and to the rest of my family and friends who never let me be discouraged. They always had confidence in my ability to finish even when I did not. Their faith gave me strength.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate School of the University of Florida in Partial Fulfillment of the Requirements of Doctor of Philosophy

THE APPLICATION OF THE STIMULUS EQUIVALENCE PARADIGM TO COGNITIVELY COMPLEX ACADEMIC TASKS AND CHILDREN WITH LIMITED ENGLISH REPERTOIRES

By

Jolenea B. Ferro

August, 1990

Chairman: William D. Wolking
Major Department: Special Education

Two experiments investigated the likelihood that equivalence classes would be formed and maintained within two cognitively complex academic tasks: reading comprehension and an English grammar task related to parts of speech. The subject in Experiment 1 could match Spanish words and pictures and was taught to match Spanish words to English words. Postassessment tests showed that he was then able to match English words to pictures (the reading comprehension task). Because of increases in baseline scores, it appeared that the subject was learning some of the relations outside of the experimental setting.

Experiment 2 was an extension of Experiment 1 using stimuli that were unfamiliar and less likely to be presented to the subjects. Two subjects were taught to match printed English words to their synonyms and to match

the same words to their parts of speech. Subsequent tests showed that one subject demonstrated the emergence of an equivalence class consisting of a word, its synonym, and its part of speech. The second subject failed to demonstrate the formation of an equivalence class, although other relations were at or near criterion levels.

The two relations that contained the part of speech deteriorated in follow-up sessions for both subjects. Subject 2 continued to exhibit baseline level responding on the tests of the equivalence relations. The subjects in both experiments were primarily Spanish speaking and were limited in their ability to use the English language. Both had been identified as exhibiting learning disabilities in their country of origin. Results are discussed in reference to naming, maintenance, test sequence, and feedback.

CHAPTER I
INTRODUCTION

Given the current emphasis on teacher accountability (Wang, 1987) and student minimum competency requirements (McCarthy, 1980), teachers need highly efficient and effective instructional procedures for the classroom. One such procedure is matching-to-sample. Matching-to-sample has been successfully used to instruct handicapped learners in both laboratory and classroom settings (Edwards, 1974; Sidman, 1971; Sidman & Stoddard, 1967). In the matching-to-sample procedure, a sample stimulus is presented, followed by a number of comparison stimuli. One comparison is then selected as a "match" for the sample stimulus. The comparison may look the same as the sample, e.g., matching the sample color brown to a comparison color brown, or it may be dissimilar, matching the sample color brown to the word brown.

The matching-to-sample procedure describes a conditional discrimination (Cumming & Berryman, 1965). That is, the relation between the sample and comparison is defined by a conditional, or if ... then statement. If A1

is presented, then choose B1; if A2 is presented, then choose B2.

Sidman, Rauzin, Lazar, Cunningham, Tailby, and Carrigan (1982) made a distinction between the conditional discrimination and the assumptions made when the procedure is called matching-to-sample. Similar stimuli are assumed to be identical whereas dissimilar stimuli are assumed to be equivalent. That is, we assume that the sample and comparison have become members of a common stimulus class. However, the typical matching-to-sample test only proves the conditional relation. Additional tests are necessary to validate the equivalence relation.

Stimulus Equivalence

Three properties define the equivalence relation: reflexivity, symmetry, and transitivity (Sidman & Tailby, 1982). All three must be present if stimuli are to be called equivalent (Sidman, Willson-Morris, & Kirk, 1986). Reflexivity is demonstrated when a stimulus is matched to itself without direct training. For example, given the sample "5", the comparison "5" is chosen from among a number of dissimilar stimuli. Reflexivity, then, corresponds to generalized identity matching.

Symmetry is demonstrated when the conditional relation is reversed. For example, the relation "if given the number 5, then choose the word five" becomes "if given the word

five, then choose the number 5." A subject who is able to make the original discrimination is required, without further training, to match the reverse sample and comparison, demonstrating functional reversibility.

In order to demonstrate transitivity, a third stimulus must be introduced. Two conditional relations are then established: "If given the number 5 then choose the word five" and "if given the word five then choose an object representation of five." Transitivity is demonstrated by the emergence of a third conditional relation in which the sample from the first relation is matched to the comparison from the second. Thus, the new relation, "if given the number 5 then choose an object representation of five", emerges. In demonstrating each property, it is important that the conditional relation is not taught or reinforced in order to be certain that equivalence is the basis for the performance. If the conditional relation is taught, the performance may be controlled by other features of the sample and comparison such as hue, brightness, or position.

The formation of equivalence classes as a product of the matching-to-sample procedure recommends it as an effective instructional model for teaching cognitively complex behaviors. In the classroom, students are frequently asked to form relationships between items that are not physically similar but are related based upon

custom, social values, or authoritatively determined rules of organization and structure. For example, students may be required to learn the stimulus class "presidents" or to categorize words according to their usage in a sentence.

In the laboratory, the formation of equivalence classes has been demonstrated with a relatively small number of basic academic skills: vocabulary and reading comprehension (Joyce & Wolking, 1989; Sidman, 1971; Sidman & Cresson, 1973), number concepts (Gast, VanBiervliet, & Spradlin, 1979), spelling skills (Mackay & Sidman, 1984), and skills relating to money (McDonagh, McIlvane, & Stoddard, 1984). Recently, Haring, Breen, and Laitinen (1989) demonstrated the formation of equivalence classes while teaching the concept of age-appropriate dress and behavior, a socially determined class of behaviors. More cognitively complex academic subject matter, i.e., subject matter such as grammar, geometry, or chemistry, that is taught late in the academic curricula, has not as yet been addressed.

The procedure also has implementation advantages that recommend it as an efficient strategy. First, instruction may be simplified into a purely visual stimulus presentation. Teaching visual classes can be delegated to electronic equipment, thus freeing the teacher for other instructional tasks. Initial reading skills such as word calling, simple reading comprehension, and spelling have

been taught using purely visual input (Mackay, 1986; Sidman & Cresson, 1973; Sidman et al., 1986).

Second, when stimulus classes are formed, i.e., stimuli become equivalent, new behaviors may emerge without having been explicitly taught or reinforced. The number of performances that must be taught directly is thereby reduced with the potential for an exponential expansion of emergent relations. For example, teaching 15 relations could generate 60 new relations (Sidman, Kirk, & Willson-Morris, 1985). This assumption is currently part of educational practice for some skills but has not been explicitly articulated as stimulus equivalence. For example, the number of multiplication facts that must be taught is thought to be limited. If a child learns 2 x 6 it is assumed that he also knows the symmetric relation 6 x 2. (Mercer & Mercer, 1981).

Third, the stimulus equivalence phenomenon allows one to take advantage of previously learned material or tasks easier to teach to individual children (Goldstein, 1985). If, for example, a child already labels certain pictures and objects, simple reading skills may be taught in one of two ways. A child taught to match printed words to pictures should be able to identify printed words that are named without further training. Conversely, a child taught to identify printed words that are named should be able to

match the printed word to the picture, simple reading comprehension.

Stimulus equivalence relations have been successfully demonstrated with both normal and mentally retarded subject populations (Lazar, Davis-Lang, & Sanchez, 1984; Mackay, 1986; McDonagh et al., 1984; Sidman, 1971; Sidman & Cresson, 1973; Sidman et al., 1986). However, the paradigm has implications for instructing other populations as well. Stimulus equivalence training may be particularly helpful in meeting the needs of the bilingual minority child.

Literature that focuses on the educational needs of the bilingual minority child calls for simplified academic input and, when possible, the inclusion of English and the minority language in training (Cummins, 1983). Bernal (1983) suggested that research efforts in bilingual education should address best practices and curriculum development. A review of stimulus equivalence research allows for the determination that stimulus equivalence is highly applicable to establishing linguistic repertoires. Goldstein (1985) proposed the use of this paradigm to maximize initial language training efforts.

The Microcomputer

When the computer was introduced, a major revolution in education was predicted. This revolution did not occur because of the relative expense, large size, and difficulty

of accessing main frame computers. Now that microcomputers are less expensive, more accessible and less complex to use, the number of computers in schools has increased dramatically (Weibe, 1987).

Desk-top computers may be particularly irresistible to educators. Many tactics identified with effective instructional practices (Englert, 1984) may be attributed to computers. For example, computer programs provide immediate feedback, control extraneous feedback, and can be personalized, i.e., rate of presentation can be controlled, the student can pace the presentation, and programs can be individualized through branching techniques. In addition, computers provide management advantages that include recording correct and incorrect responses, recording latencies between stimulus presentations and response, and charting progress.

Computer use for individuals with special needs has been enthusiastically supported because of its potential for providing more effective, independent, and functional experiences (Behrmann & Lahm, 1984; Garner & Campbell, 1987; Hannaford, 1987). Some experimenters have concluded that computer use results in increased attention, motivation, and time on task (Hannaford, 1987). Students seem to learn the material in less time and computers have been helpful in reviewing previously learned materials (Hannaford, 1987).

However, there are many obstacles and cautions that apply to computer use in special education. First, more research is necessary to verify satisfactory and meaningful improvement in the performance of handicapped students (Behrmann et al., 1984; Hofmeister, 1982; Tawney & Sniezek, 1985). Second, the quality and quantity of software is inadequate (Billings, 1983). Software must be developed that effectively enhances learning rather than serving the function of an elaborate page turner. Third, computer software must be designed that can be adapted by the teacher to present current classroom material.

Procedures designed to generate and test equivalent relations are adaptable to computer presentation. However, computer applications that may be implemented in the classroom have not as yet been devised. It is not enough to suggest that teachers can read the literature and make adaptations needed for their classrooms. Teachers responses toward electronic equipment are varied: They can fear the technology, they can lack computer skills that would allow them to program a specific technique or procedure, and they can have little time to implement and evaluate particular programs.

One solution is to design a computer program that has a basic instructional procedure and into which the teacher may enter specific curricula. Many authoring systems provide

this capability. The user of one of these systems is only required to enter specific information and does not have to learn to program the computer in order to create individualized instruction.

Statement of the Problem

Although stimulus equivalence training has the potential for application to a variety of academic tasks and a number of student populations, research addressing its applicability to more cognitively complex academic subject matter has been limited. Instead, the research directly related to academic subject matter has been conducted in the laboratory and has been directed toward the formation of equivalence relations in basic academic tasks, such as money concepts.

Classroom applications of the laboratory research have not generally been made. One reason for this may be that laboratory demonstrations have often used complicated procedures or have used both complicated procedures and elaborate equipment. Advances in microcomputer technology and authoring systems provide the vehicles for implementation of new and innovative instructional techniques. The stimulus equivalence paradigm seems particularly suited to microcomputer application in that stimulus relations have been taught without the need for teacher intervention. However, no computer programs that

use the stimulus equivalence phenomena to enhance instruction are currently available to teachers.

Another limiting factor is that the special needs population that has been most studied has been the mentally retarded. However, the paradigm has the potential for use over a much broader spectrum of subjects and disabilities. Bilingual students have special needs relating to language performance that may be addressed using stimulus equivalence procedures.

Although it has been recommended as an effective teaching technique, stimulus equivalence is relatively unknown and unused in practice. Research that may encourage wider educational use is needed.

This study is important for several reasons. First, it will extend stimulus equivalence research to include more cognitively complex academic subject matter. Experiment 1 investigated whether previously learned relationships between Spanish words and pictures could be used to establish a purely visual equivalence class consisting of Spanish words, English words and pictures. Experiment 2 investigated whether an equivalence class could be formed between words and their parts of speech, i.e., noun, adverb, adjective, or verb. Second, this study will provide data relevant to the maintenance of the classes generated in Experiment 2.

Third, it will provide data that should allow for the determination of the effectiveness of stimulus equivalence training for bilingual students. A review of the recent literature has enabled this investigator to conclude that it would be advantageous to develop appropriate and effective instructional technology for the bilingual population.

Fourth, this study will provide information related to the use of the microcomputer with special needs students. There is very little empirical support for much of the information presented about computer use with special needs students. Fifth, classroom applications of stimulus equivalence training and testing have been limited. This study has the potential to provide the means for more extensive classroom use.

Questions Under Investigation

This study will investigate whether limited English speaking students will form equivalence classes on two cognitively complex academic tasks. First, will limited English speaking subjects, who are able to match Spanish words to pictures, demonstrate the ability to match English words to pictures (reading comprehension) after they are taught to match Spanish words to English words? Second, will limited English speaking subjects form an equivalence class consisting of an English word, its English synonym, and the word that identifies the correct part of speech when

they are taught to match a word to its synonym and to match a word to its part of speech. Third, will the performance on the second task be maintained after two- and four-week intervals.

Delimitations of the Study

This study was delimited in several ways. The investigation was conducted in Alachua County, a medium-sized county located in the north central portion of Florida. Subjects for the study were bilingual students whose primary means of communication was Spanish. They were selected from one middle school in Gainesville. Subject participation was limited to students between 11 and 14 years. No consideration was given to sex or social economic status.

Limitations of the Study

Results of this study should not be generalized to other bilingual populations. In addition, they should not be generalized to other handicapped or nonhandicapped students. Generalization to other academic tasks should not be made without systematic replications.

Summary

Recent research into stimulus class formation has identified procedures that capitalize on the development of equivalence relations. These procedures can result in the emergence of large numbers of previously unreinforced and

untrained new relations, thus reducing the number of performances that must be directly taught. The procedures are suited for microcomputer applications and may be effective across a variety of student populations. However, research has not addressed a variety of complex academic tasks and the procedures have not been implemented in the classroom. This study was designed to investigate emergence of equivalence classes of complex academic tasks using bilingual students as subjects. The procedures were implemented using an authoring system that has the potential to be easily and efficiently applied in the classroom.

A review of pertinent research related to stimulus equivalence is presented in Chapter II. Chapter III presents the methodology used in this investigation. The results and their implications are discussed in Chapter IV and V respectively. Definitions of terms used in this study are presented in the Glossary.

CHAPTER II
REVIEW OF THE LITERATURE

A large part of the research on stimulus equivalence has been conducted in the laboratory and has been directed toward the identification of the manner in which equivalence relations are formed. Most studies use nonsense symbols, words, and objects to avoid problems associated with pretesting and the possibility of learning outside of the experimental situation. Only a few studies have directly addressed the acquisition of cognitively complex behaviors using meaningful stimuli. The formation of equivalence classes has been demonstrated in reading (Sidman, 1971), premathematic skills (Gast et al., 1979), spelling skills (Mackay & Sidman, 1984), skills relating to the use of money (McDonagh et al., 1984; Stoddard, Brown, Hurlbert, Manoli & McIlvane, 1989) and daily living skills (Haring et al., 1989).

Only one study to date has evaluated the formation of stimulus classes in the acquisition of a second language (Joyce, Joyce, & Wellington, 1990). However, stimulus

equivalence research in which manual signs (VanBiervliet, 1977) or Greek letters and words (Sidman & Tailby, 1982) were used as stimuli has some relevance to this issue.

The following review of the stimulus equivalence research includes a historical overview, an analysis of studies in which meaningful stimuli were used and that may be pertinent to the acquisition of cognitively complex academic skills, and an analysis of studies that included stimuli in a language other than English (Spanish words, manual signs, Greek and Hebrew words and letters).

Stimulus Equivalence

Overview of Stimulus Equivalence Research

Since 1971, research has demonstrated that equivalent stimulus classes, consisting of arbitrary stimuli, could be developed within a matching-to-sample task (Mackay, 1986; Sidman, 1971; Sidman & Cresson, 1973; Sidman, Cresson & Willson-Morris, 1974; Sidman & Tailby, 1982; Spradlin, Cotter, & Baxley, 1973). Stimuli are arbitrary, or nonidentical, when they are physically dissimilar from one another. Much of language relies on equivalence among stimuli. That is, a stimulus must evoke a response that is physically dissimilar from the stimulus. For example, a picture of a dog, the spoken word 'dog', and the printed word "dog" are physically dissimilar stimuli that may be members of the same stimulus class.

Sidman et al. (1982) noted that when 'arbitrary' stimuli are matched, investigators assume that an equivalence relation has been formed. However, matching, or simply applying the same name to the stimuli, is not sufficient evidence to establish that an equivalence relation exists (Mackay & Sidman, 1984). Rather, all of the three following relations must be verified to demonstrate that an equivalence class has been formed: a) reflexivity, or generalized identify matching; b) symmetry, or reversibility of sample and comparison; and c) transitivity, or the appropriate recombination of samples and comparisons into a new conditional discrimination (Sidman & Tailby, 1982).

In theory, the three relations are hierarchical, i.e., reflexivity precedes symmetry, which precedes transitivity (Spradlin & Saunders, 1984). In fact, the hierarchical structure of the relations has not been proved. However, the empirical evidence seems to be congruent with this interpretation. First, when stimulus class development has not been demonstrated (subjects have failed to display transitive relations), further testing has often revealed that symmetrical relations have also not been formed. Once the symmetrical relations are learned, transitive relations also emerge (Spradlin & Saunders, 1986).

Second, stimulus class formation may be enhanced when tests of symmetry are conducted prior to tests of transitivity. Currently, many investigators train two conditional relations (e.g., match the number "5" to the word five and match the number "5" to five objects) and then test for equivalence using a combined test that permits the simultaneous evaluation of symmetry and transitivity (see Sidman & Tailby, 1982, for a detailed description of this combined test). If the equivalence relation is not demonstrated, i.e., the subject fails the combined test, the trained relations are retested and the symmetrical relations are presented in a test that is not combined with the test for transitivity. There has been a great deal of variability among subjects in the number of tests and the time required for the formation of an equivalence class.

Stoddard and McIlvane (1986) suggested that the successful emergence of equivalence relations might be more readily produced by evaluating the three relations directly and in their hierarchical sequence. Recently Haywood, Fields, Adams and Verhave (1990) presented preliminary data addressing the issue of test order and the formation of equivalence classes. They compared two sequences of testing. In one sequence, complex to simple, the equivalence relation was tested first, as is the most common practice. Tests of other relations were conducted only if equivalence class development was not demonstrated. In the

second sequence, simple to complex, the symmetrical relation was tested immediately after the subject met criterion on the trained relation (e.g., train A-B and immediately test B-A). All trained and symmetrical relations were demonstrated before the test for the transitive relation was administered. Last, the transitive relation and its symmetrical counterpart were evaluated rather than combining the equivalence test.

The authors found that all subjects who were tested in the simple to complex sequence formed equivalence relations faster than subjects tested in the complex to simple sequence. In addition, intersubject variability in the amount of time required to form the equivalence relation was reduced for the subjects who received the simple to complex sequence.

When stimuli are shown to be equivalent, they are commutable, i.e., one may be substituted for another (Goldstein, 1985; Spradlin & Saunders, 1986). This substitutability allows for class expansion by relating a new stimulus to one of the existing class members. The new stimulus can then become equivalent to all class members. All that is known about each of the existing class members may then be acquired by the new stimulus.

Spradlin et al. (1973) used nonsense forms to test the possibility of establishing a four-member class by directly relating the fourth stimulus to only one member of a three-

member class. Six mentally retarded adolescents participated, three in each of two experiments. In the first experiment, subjects received pretraining to match auditory words (_A_) to pictures (_B_) in a two-choice discrimination. Subjects were then trained to match auditory words (_A_) to a printed word (_C_) and to match pictures (_B_) to printed words (_C_). Finally, auditory words (_A_) were matched to oral naming (_D_). When the subjects were tested for the emergence of the B/D relation (matching the picture to oral naming), two demonstrated the formation of a four-member class. One required repeated testing, but no additional training, to demonstrate class formation. In experiment two, pretraining was eliminated. Three subjects demonstrated that a four-member class had been established. In these two experiments, pictures (_B_) were indirectly related to oral naming (_D_) through printed words (_C_).

Unfortunately, Spradlin et al. (1973) neglected to test for all relations, as is necessary to demonstrate stimulus equivalence. Sidman and Tailby (1982) used Greek letters to teach four-member stimulus classes to 8 normal children in a manner similar to that described by Spradlin et al. In this experiment all possible relations were tested. Six children responded in a manner that indicated the formation of stimulus classes containing four members. Testing the two remaining children indicated that some prerequisite symmetric relations had not been formed.

Subsequent investigations with normal children have demonstrated expansion of class membership to five, using purely visual stimuli (Lazar et al., 1984), and expansion to six members when subjects were taught prerequisite skills for classes of gradually increasing size (Sidman et al., 1985).

Saunders, Saunders, Kirby, and Spradlin (1988) have demonstrated the formation of 9-member classes in mentally retarded subjects. They noted the possibility that the addition of more members to a class may increase the stability of the relations within the class making it less susceptible to disturbance. For example, if one linking relation in a three-member class is broken, the entire class may disintegrate. However, if one linking relation in a nine-member class is broken, the other linking relations may continue to maintain the class.

The addition of initially unrelated stimuli to an existing stimulus class also has implications for generalization. It allows information learned in one time and place to be related to information encountered in another time and place (Wulz & Hollis, 1979). Thus, stimulus equivalence may help explain the conditions under which an individual can behave appropriately in a new situation or to stimuli never before encountered. If this were the case, the network of linking relationships generated through stimulus equivalence could result in more

meaningful, better maintained, and more easily generalizable skills.

One of the first studies to actually test generalization of equivalence classes was presented by Reeve, Fields, Adams, Kaplan, and Goss (1990). The authors tested generalization of equivalence relations using stimuli that consisted of different length lines. They found that class membership did generalize to lines of similar length.

Equivalence Studies Pertinent to the Acquisition of Academic Subject Matter

Early tests of the stimulus equivalence paradigm demonstrated reading comprehension in mentally retarded persons (Sidman, 1971; Sidman & Cresson, 1973; Sidman et al., 1974). Sidman and Cresson (1973) conducted an experiment with two severely retarded boys that was a replication of an experiment conducted by Sidman in 1971. The purpose of the experiment was to demonstrate a procedure that would facilitate the transfer of auditory-visual equivalences (auditory comprehension in which auditory stimuli were matched to visual stimuli or vice versa) to visual-visual equivalences (simple reading comprehension in which stimuli were purely visual, i.e., pictures & words). Tasks that require a matching performance may be used as tests of simple reading comprehension if an equivalence class is formed (Sidman et al., 1982). The formation of equivalence classes allows for the interpretation that the

picture and the printed word have the same meaning (Sidman, 1988).

Four stimuli, A, B, C, and D, were included in each class. Stimuli in set _A_ consisted of auditory words, set _B_ stimuli consisted of pictures, set _C_ stimuli consisted of printed words, and set _D_ stimuli consisted of oral naming. Figure 2-1 is a schematic representation of the stimulus sets. Solid lines indicate relations that were trained, and broken lines indicate relations that were tested.

A test battery, administered to the subjects, contained the following tests: a) matching a printed word to a printed word (identity matching, C/C); b) selecting a picture to a dictated word sample (auditory comprehension, A/B); c) orally naming a picture, B/D); d) selecting a printed word to a dictated word sample (auditory receptive reading, A/C); e) matching a printed word to a picture (reading comprehension, B/C); f) matching a picture to a printed word (reading comprehension, C/B); and g) naming the printed word samples (oral naming,C/D). Seven forms of the test battery were used to test the subjects prior to training, after each training sequence, and when all training had been completed. The training sequence was: a) teach identity matching; b) teach auditory comprehension, relation A/B; c) teach auditory receptive reading with 9 words, relation A/C; d) teach 5 more words (14); and e) teach all words (20).

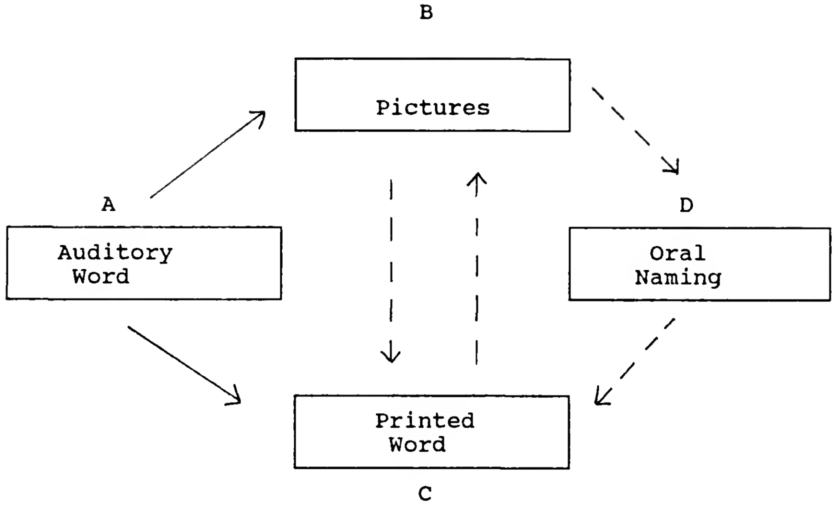


Figure 2-1. Schematic representation of Equivalence Relations demonstrated by Sidman and Cresson, 1973.

Note. In the diagram, solid arrows designate relations known or taught during the experimental phase. Broken arrows designate equivalence relations tested.

Two subjects demonstrated reading comprehension skills, relations B/C and C/B, commensurate with the number of words that were taught in the auditory receptive task. That is, when auditory receptive reading was taught using 9 words, subjects were able to correctly match 9 picture samples to the corresponding printed words. When all 20 words had been taught, subjects were able to perform both reading comprehension tasks and to name the printed words aloud.

This study was important for several reasons. First, the authors successfully established a relationship between an auditory receptive task, a purely visual task and a verbal production task. They did this without directly reinforcing either the visual task or the verbal production task. Second, reading skills were taught using an automated display box, touch windows, and a tape recorder. Thus, the time required for teacher intervention was minimized. Third, the successful application of the stimulus equivalence paradigm stimulated research activity that has implications for the study of semantic processes, the development of language, and the development of cognitively complex academic skills.

However, this study also generated a number of unanswered questions. First, the role played by the oral naming task was not clear. Was oral naming a necessary component of stimulus class formation or was it a by-product? Second, stimulus classes included auditory as well

as visual stimuli. It was not clear whether purely visual stimulus classes could be formed. Third, although both subjects learned the comprehension and oral naming tasks, one subject required two months of teaching and the second subject required 11 months of teaching. In addition, the authors indicated that they were required to change the teaching procedures several times to overcome 'difficulties'. It is not clear in what manner changing procedures influenced the stimulus class formations.

Sidman et al. (1974) explored the role of naming in the formation of stimulus classes and tested the generality of earlier findings using lower and upper case alphabet letters. Subjects in the Sidman and Cresson (1973) experiment were taught to match dictated words to both pictures and printed words. The subjects described by Sidman et al. were only taught to match dictated words to upper-case letters. The dictated names of lower-case letters were never taught. Thus, the possibility that subjects in the previous experiment learned to read words aloud when they were taught to match printed words to dictated names was eliminated with the new procedure.

The authors taught two subjects to select upper case letters when they were named, to match upper case letters to lower case letters, and to match lower case letters to upper case letters. A test battery evaluating all stimulus relations was administered prior to training, after the

completion of each teaching sequence, and after all stimuli had been taught.

The mentally retarded subjects in this experiment were able to form equivalence relations without direct training between lower case letters and dictated names. However, the subjects demonstrated limited ability to read the letters orally. The authors suggested that oral naming may not be necessary for the emergence of visual-auditory matches.

Although the authors extended earlier findings, many questions were still unanswered. The role of naming had not been sufficiently clarified. The manner in which a subject is asked to name a stimulus, or some other unrecorded aspect of the procedure, could have been responsible for the subject's inability to read the letters orally. The subjects in this experiment also required an extended period (4 months) of frequent training to learn the trained equivalence relations. The question of purely visual matches had also not been resolved.

Furthermore, no procedures were implemented to verify accurate data collection in the experiments described above (Sidman, 1971; Sidman & Cresson, 1973; Sidman, et al., 1974). This was true whether the data were recorded manually by the experimenter; for example during the oral naming tasks, or were recorded automatically by the stimulus presentation apparatus, during all other tasks. In addition, teaching procedures were not standardized.

Thus, stimulus class formation may have been due to some unknown training variable.

In spite of the design problems, the authors cited above demonstrated with mentally retarded subjects that a picture, a spoken word, and a printed word can become equivalent members of the same class. Each demonstration of the formation of a stimulus class consisting of a dictated word or letter, picture, printed word or letter, and named word or letter is a direct replication of the development of stimulus equivalence relations. In addition, tests administered after teaching some relations demonstrated that untrained relations did not emerge before the appropriate training had been accomplished. Joyce and Wolking (1989) also demonstrated the emergence of reading skills (oral naming and comprehension) using preschool subjects.

McDonagh et al. (1984) systematically replicated and extended the stimulus equivalence paradigm to the formation of purely visual stimulus classes. Although purely visual classes had been demonstrated prior to this study, visual stimuli had consisted of arbitrary forms with no practical application (Spradlin et al., 1973). In the McDonagh et al. study, stimuli included groups of coins and printed prices. Oral naming was tested but was never taught, and the subject was never asked to select a stimulus when the sample was a dictated word.

The subject, a mentally retarded female, was first evaluated on the simple three-member class represented by the schematic diagram in Figure 2-2. The solid lines represent relations that were taught and the numbers designate the sequence. After two training sessions, the subject demonstrated the formation of a stimulus class consisting of the printed price (5 cents), 5 pennies, and a nickel.

In the second set of relations represented in Figure 2-3, the subject matched 10 pennies to a dime, after having been taught to match the printed price (10 cents) to a dime and to match the printed price (10 cents) to 10 pennies. The subject also matched 2 nickels to 10 pennies, after having been taught to match the printed price (5 cents 5 cents) to 10 pennies and to match the printed price (5 cents 5 cents) to two nickels. However, the subject was initially unable to successfully exhibit the stimulus relations represented by the sequence numbers 10 through 18 in Figure 2-3. This subject failed to demonstrate equivalence in a larger class until more relations were trained and tested using a procedure the authors identified as constructing responses within a matching-to-sample task. Constructed response matching-to-sample consisted of selecting coins or prices that matched the sample displayed on an index card from an unsorted pool of coins.

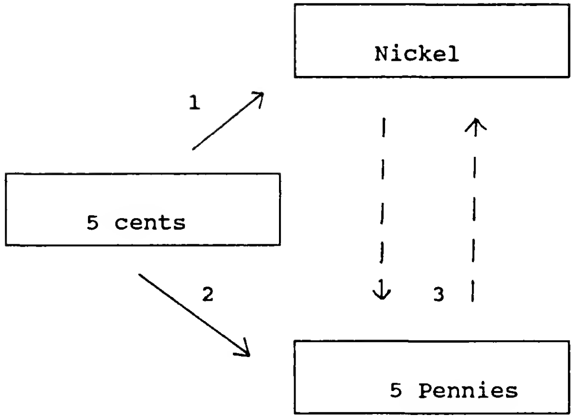


Figure 2-2. Schematic representation of Equivalence Relations demonstrated by McDonagh et al., 1984.

Note. In the diagram, 5 cents is a printed price. Solid arrows designate relations known or taught during the experimental phase. Broken arrows designate equivalence relations tested. Numbers designate the sequence of teaching and testing.

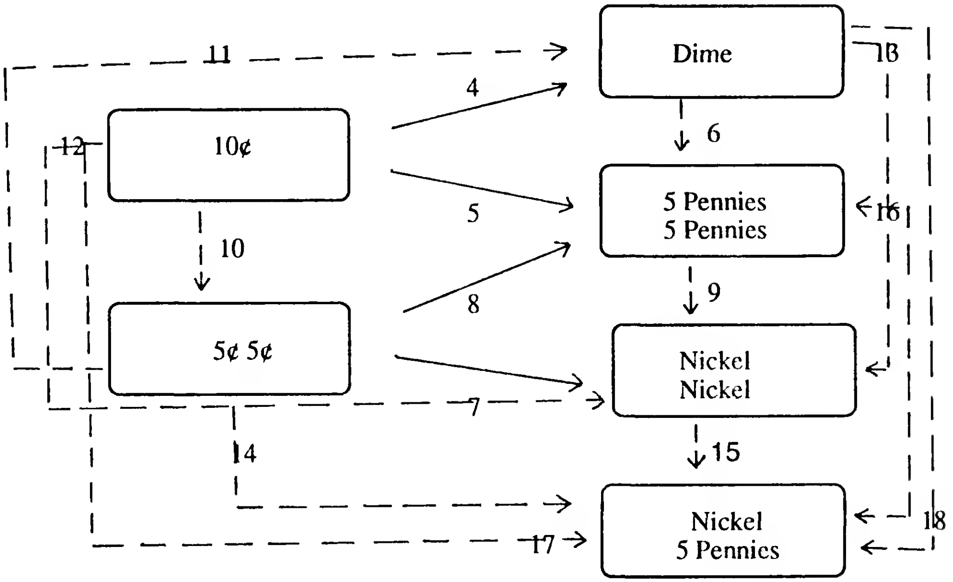


Figure 2-3. Schematic representation of Second Set of Equivalence Relations demonstrated by McDonagh et al., 1984.

Note. In the diagram, 5 cents and 10 cents are printed prices. Solid arrows designate relations known or taught during the experimental phase. Broken arrows designate equivalence relations tested. Numbers designate the sequence of teaching and testing.

Stoddard, et al. (1989) continued the study of monetary equivalences with three mentally retarded subjects and coin combinations up to 50 cents. The authors used a combination of standard matching-to-sample, exclusion training and component matching to teach the required relations. Component matching consisted of breaking down combinations into lower value components, e.g., one component of 20 cents is two dimes. Some components were taught and others were tested to see if they emerged. Testing and training consisted of constructed response matching-to-sample. Using this procedure, all subjects demonstrated the emergence of new relations.

Failure to demonstrate the larger class of equivalent stimuli in the coin study may have been the result of too many possible new relations, the training and testing procedure, or flaws in the design of both studies. First, no attempt was made to assess the accuracy with which the procedure was implemented. It is likely that stimulus presentation varied across a number of dimensions. Stimuli were manually placed on a display panel that was laid flat on a table. The author did not indicate that precautions were taken to ensure standardized display, a standard intertrial interval, or inadvertent interaction with the subject. Second, no reliability data were presented that would indicate accurate data collection.

Third, training trials, test trials, baseline and review trials were presented together during testing procedures. Correct responses during training trials were reinforced but correct responses during test trials were not. Differences in reinforcement procedure may lead to a discrimination between the two types of trials and to differential responding (Bush, Sidman & deRose, 1989).

The value of these two studies is that the subjects demonstrated the formation of a stimulus class using purely visual stimuli. Mackay (1986) also demonstrated purely visual stimulus class formation with anagram naming (spelling). These findings suggests that subjects may be able to perform reading comprehension tasks in the absence of experimenter provided labels.

Equivalence Studies that Included Stimuli in a Language Other than English

Two studies demonstrated the formation of equivalence classes with a sign language component (VanBiervliet, 1977; Wulz & Hollis, 1979). Although not spoken, manual signs may be a second method of communication after an individual's native language. As such, acquisition of manual signs within the stimulus equivalence paradigm is directly relevant to this investigation.

Investigators in the studies cited in the paragraph above were interested in techniques that might enhance language acquisition in individuals with a history of

minimal language reception and production. The stimulus equivalence paradigm was chosen because it would seem to predict that sign-word and sign-object training would result in word-object association skills, a necessary prerequisite for many language acquisition programs (VanBiervliet, 1977, p. 178).

VanBiervliet included 3 members in each stimulus class: nonsense words, junk objects, and nonsense manual signs. He used nonsense rather than real symbols to avoid the time required for pretesting. Six mentally retarded subjects participated. Each subject began the experiment with some receptive and productive language skills. Prior to training, subjects were randomly divided into three pairs. Each subject pair was trained to imitate a sign, produce a sign in response to an object, point to an object in response to a sign, imitate a word, produce a sign in response to a dictated word, and say a word in response to a sign. Object identity matching was trained when necessary. Figure 2-4 presents a schematic representation of the training and testing relations.

Following training, the formation of stimulus classes was assessed by testing word-object reception (point to an object in response to a dictated word) and object-word production (say a word in response to an object). Test sessions were separate from training sessions, and no reinforcement was provided during testing. The results

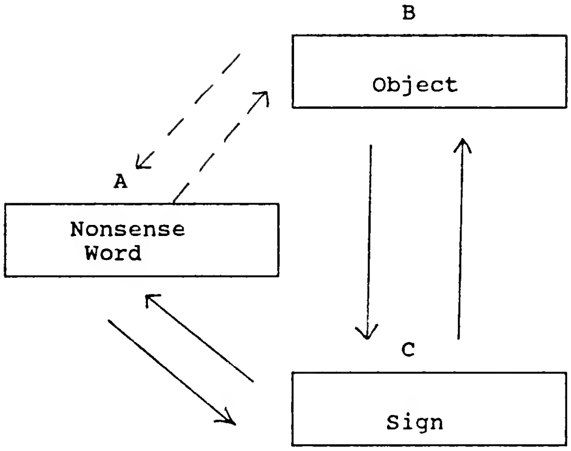


Figure 2-4. Schematic representation of Equivalence Relations demonstrated by VanBiervliet, 1977.

Note. In the diagram, solid arrows designate relations known or taught during the experimental phase. Broken arrows designate equivalence relations that were only tested.

indicated that all subjects were able to successfully perform both word-object tasks.

Wulz and Hollis (1979) demonstrated the formation of stimulus classes with five mentally retarded subjects using abstract pictures, printed nonsense words, printed nonsense symbols, dictated words, and nonsense signs. Subsequent to training, subjects had successfully formed stimulus classes that included a picture, a printed word, a dictated word, a symbol and a sign.

Recently Joyce et al. (1990) presented data demonstrating the formation of equivalence classes that included Spanish words. Their subjects were two adolescents diagnosed with traumatic brain injury. Preassessments determined that both subjects could name orally and match English printed words used as stimuli. Subjects were trained to match Spanish printed words and pictures. Post assessment data showed the emergence of symmetrical and equivalence relations. Subjects were able to match Spanish words and English words, Spanish words and pictures, and to translate dictated English words to Spanish words and dictated Spanish words to English words.

The experiments cited above included auditory as well as visual stimuli in the stimulus classes. Other investigators have used languages unfamiliar to the subjects as stimuli in evaluations of stimulus class formation. Greek and Hebrew words and Greek letters were used to avoid

lengthy pretesting (Lazar et al., 1984; Sidman et al., 1985; Sidman et al., 1986; Sidman & Tailby, 1982). Acquisition of the unfamiliar language itself was not relevant to the purpose of these studies.

One of these studies (Lazar et al., 1984) is relevant to this investigation because it evaluated the enlargement of class membership with matching tasks involving purely visual stimuli. Stimuli were 15 upper and lower case Greek letters and script Hebrew letters. Four normal children were taught a series of conditional relations; each trained relation had the potential to enlarge the stimulus class until it contained five members. Tests for the emergence of stimulus classes were implemented after each training series. Equivalence tests were interspersed with training trials but no reinforcement was available for any correct answer. Three subjects formed all equivalence relations. The fourth subject demonstrated stimulus class formation after the testing procedure had been altered to include only a limited number of relations within one test session.

Summary

Previous studies have demonstrated the emergence of stimulus classes that are relevant to the present investigation. First, experimenters have established large classes by relating a new stimulus to one member of the existing class (Lazar et al., 1984; Saunders et al., 1988;

Sidman & Tailby, 1982; Sidman et al., 1985; & Spradlin et al., 1973).

Second, some cognitively complex academic tasks have been evaluated in previous investigations. Sidman and his colleagues (Sidman, 1971; Sidman & Cresson, 1973; Sidman et al., 1974) and Joyce and Wolking (1989) described the formation of equivalence classes that demonstrated simple reading comprehension. Word-picture and picture-word matches emerged after subjects had been trained to match pictures to dictated words and printed words to dictated words. The stimulus equivalence paradigm has also been extended to number concepts (Gast et al., 1979), spelling skills (Mackay & Sidman, 1984) money skills (McDonagh et al., 1984) and the concept of age appropriateness (Haring et al., 1989).

Third, although many studies have evaluated cross modal stimulus class formations, purely visual stimulus equivalences have also been demonstrated. The success of visual-visual training makes these procedures highly adaptable to computer presentation (Stoddard & McIlvane, 1986). Spradlin et al. (1973) and Wetherby et al. (1983) demonstrated purely visual class formation with forms as the stimuli. Lazar et al. (1984) more nearly resembled the investigation proposed in this paper by using Greek and Hebrew letters in a purely visual stimulus equivalence test.

Fourth, subjects have been able to form stimulus

classes that included other languages (Joyce et al., 1990; VanBiervliet, 1977; Wulz & Hollis, 1979). The manual signs used in two studies cited above are a method of communication. As such, they may be referred to as a language.

This study was designed to extend earlier findings to a new subject population and to more cognitively complex academic subject matter. In addition, maintenance of all relations was tested.

Subjects were limited English speaking students in a bilingual classroom. In experiment 1, the emergence of 3-member classes of purely visual stimuli containing a Spanish word, a picture, and an English word was evaluated. New stimuli (English words) were added to an already existing stimulus relationship between pictures and Spanish words. Equivalent class formation was verified through performance of a reading comprehension task (picture-word match) that required the translation of an English word.

In Experiment 2, the research evaluated the formation and maintenance of equivalence relations in a cognitively complex academic subject area, English grammar.

CHAPTER III
METHOD

This study extended the stimulus equivalence research to a new subject population and to more cognitively complex academic tasks. The study had three purposes. First, it was designed to determine if students with limited proficiency in the English language, who could match Spanish words to pictures, would correctly match English words to pictures after they had been trained to match Spanish words to English words. Second, the study was designed to determine whether students from the same population would form an equivalence class consisting of an English word, its English synonym, and the word that identifies the correct part of speech when they were taught to match a word to its synonym and the same word to its part of speech. Third, the study was designed to test maintenance of the second equivalence relationship at two and four week follow-up sessions.

In addition, this study provided an opportunity to gather information related to the use of a computer application in the classroom. Stimulus equivalence

procedures seem particularly well suited for microcomputer application. If the procedure demonstrates improvement in student performance, students indicate an eagerness to work at the computer, and the specific curriculum content can easily be modified, the computer application is a potentially valuable instructional tool. This chapter describes the subjects, apparatus, setting and experimental procedures used in this investigation.

Experiment 1 was a pilot study that provided data relevant to the formation of a purely visual equivalence class. The emergent relation tested in Experiment 1, matching a picture to an English word, has been identified as a reading comprehension task (Sidman & Cresson, 1973). In this study, the task was assumed to be mediated visually through the previously learned relationship between the Spanish word and the picture. Experiment 1 consisted of six phases: stimulus selection, baseline, two training phases, midassessment and postassessment.

Experiment 2 provided data relevant to the formation of purely visual equivalence classes within a cognitively complex content area of instruction. Experiment 2 consisted of five phases: baseline, training the A-B relationship, training the A-C relationship, postassessment, and maintenance.

Permission for this study was obtained from Alachua County School administration, The University of Florida Institutional Review Board (Committee for the Protection of Human Subjects), and subject's parents (See Appendix A for a copy of all permission forms).

Experiment 1

Subjects

One subject, a 12 year 10 month old male from Puerto Rico, participated in this study. He had been enrolled in an 8th grade middle school program for limited English speaking students in Alachua County for two months. He spent three class periods a day in the bilingual classroom. Although the subject knew some English vocabulary, his primary method of communication was Spanish.

Upon his admission to the middle school program, school personnel administered two English Language tests, the Dade County Aural Language Survey and the Arlington Oral Language Test. He had not passed any item on these two tests. These tests are designed to assess a student's receptive and expressive use of the English language. Failure to pass an item is used as an indication that the student has no, or extremely limited, understanding of English.

The Stanford Binet was administered in 1983 in Puerto Rico and this subject attained an I.Q. score of 110. At that time, he was also diagnosed as having problems with

long term memory, visual/motor coordination, body balance, and pencil grasp. Throughout his schooling he had difficulty with reading and writing skills. Intervention was provided through special school programs and tutoring. However, the only special service he was receiving in the United States was bilingual training.

Apparatus and Setting

Sessions were conducted 5 days per week during school hours in a quiet area of the classroom. The subject was seated at a table facing a Macintosh SE computer equipped with a hard disk and mouse. The experimenter was positioned behind and slightly to one side of the subject during each session. During testing, three sessions per day were scheduled; one during each class period. During training, only one session was scheduled each day.

Stimuli were displayed on a monochrome monitor screen (19 x 14 cm). Stimuli were presented within transparent rectangular windows, measuring 5 cm by 2 cm. A sample window was positioned in the center of the monitor screen. Six windows for comparison stimuli surrounded the center window. Only three comparison windows were visible during any one experimental trial. The three comparison stimuli were positioned in one of two configurations that occurred on alternate stimulus presentations. Each configuration presented the comparison windows in a triangular

arrangement. That is, two comparison stimuli were positioned on either side of the sample, and one comparison stimulus was positioned either above of below the sample (see Figure 3-1). In each configuration, the transparent windows were 2 cm from the center stimulus window.

The cursor on the monitor screen was represented in the shape of a small hand, measuring 2 cm. The hand could be moved to any position on the screen by moving the mouse across a pad placed on the table. When the hand was positioned on a stimulus, pushing a button on the mouse resulted in a variety of programmed consequences described below.

A computer program controlled stimulus presentations during all phases and automatically recorded data during the experimental and maintenance phases. During oral naming tasks, verbal responses were recorded by the experimenter. For all other trials, the computer recorded the number of responses (correct and error responses recorded separately) and session length. A summary of each session's data was automatically recorded in individual subject files by the computer. Data were transferred by the investigator from individual subject files to spreadsheets and charting programs for analysis.

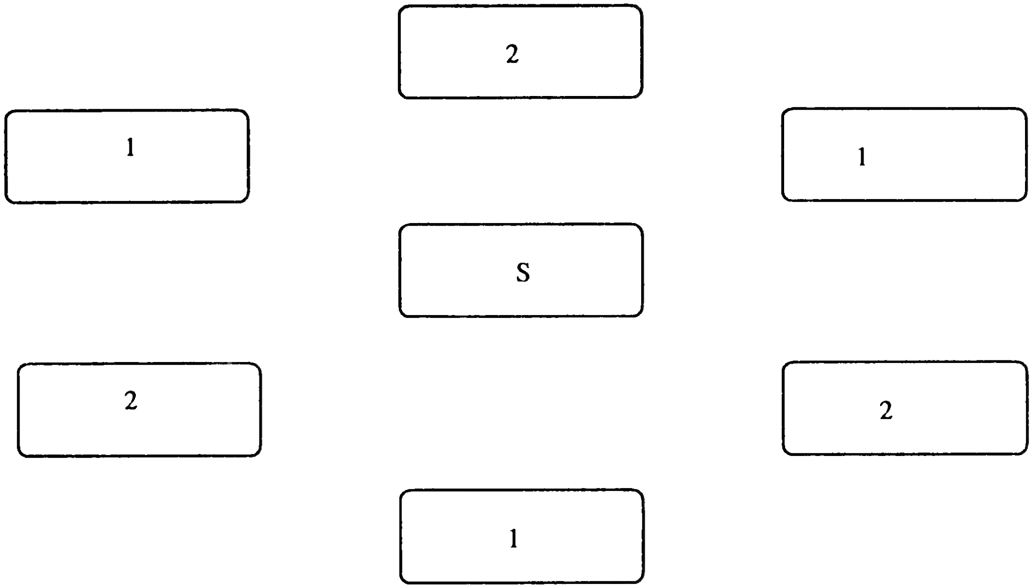


Figure 3-1. Schematic representation of Window Configuration used in Experiment 1 & 2.

Note. S designates the sample window. Windows numbered 1 and 2 were used as Configuration 1 and 2 respectively in Experiment 1. Comparisons were randomly presented in four windows in Experiment 2.

The computer program was written using the authoring environment of HyperCard (Apple Computer, Inc.). This system was chosen in order to facilitate later use of the program by classroom teachers. HyperCard provides a number of tools, such as buttons, icons, and sounds, and also permits the use of Apple's high resolution graphics for picture presentation.

Experimental Procedure

Stimulus presentation

To begin each session, a sample stimulus was displayed in the center of the screen. During oral naming tasks, the subject's response (saying a word or "I don't know") was followed by the experimenter placing the cursor on the sample window and pressing the mouse button. This initiated a 2 second (s) intertrial interval (ITI) in which the stimulus was removed from the screen. At the end of the ITI, a new stimulus was presented.

During all other training or testing procedures, placing the cursor on the sample and pressing the mouse button resulted in the display of comparison stimuli concurrent with the sample. Further presses on the center stimulus had no programmed consequences. Any presses of the mouse button that occurred while the hand was outside the borders of a transparent window were not recorded by the computer and had no programmed consequences. The order of

presentation of sample stimuli was programmed to occur randomly, as were the distractor words and the position of the correct comparison stimuli.

Consequences for correct and error responses

Test trials. Correct and error responses were consequated in the same manner for all test trials. That is, an ITI was programmed in which stimuli were removed from the screen for 2 s followed by presentation of the next sample.

Training trials. When a correct response was made, the correct word was highlighted, a chime sounded, all stimuli were removed from the screen, and the next sample was presented. When an incorrect response was made, a two step procedure was implemented. First, the incorrect word was highlighted, the computer made a "boing" sound, the correct response was underlined and an arrow pointed to it. The subject was required to select the correct match with these cues present. During step one, presses on the incorrect comparisons or the sample were counted but no change in the visual stimuli on the screen was programmed. When the subject pressed on the correct comparison, the stimulus was highlighted, a chime sounded and the second step of instruction was presented.

During step two, the same sample and comparison stimuli were immediately presented with comparison stimuli in the alternate configuration. The subject was now required to select the comparison stimulus designated as correct without the aid of the training cues programmed in step one. Presses on the sample or incorrect comparisons were again counted, but no change in the visual stimuli on the screen was programmed. Pressing the correct comparison resulted in the correct word being highlighted, a chime sounding, all stimuli being removed from the screen, and presentation of the next sample, beginning a new training trial.

Stimulus selection and baseline procedures

Stimulus selection. Before the experiment began, two lists, each containing 200 possible stimulus words, were developed. One list contained Spanish words, and the second list contained English translations. These lists were divided into sets of 25 words each to be used during the stimulus selection phase.

Words were nouns chosen from the English for Speakers of Other Languages (ESOL) Curriculum Guide provided by the Bilingual Education classroom teachers. Synonyms and hyphenated words were not included in the list. Cognates, English and Spanish words that had a common root, were also not included. An effort was made to identify Spanish words that were familiar to the subject by talking to the

bilingual classroom assistant who was also from Puerto Rico. This list of English and Spanish words is contained in Appendix B.

Figure 3-2 shows the three stimulus sets used during the experiment. Stimulus selection was based upon the results of one oral naming task and two picture/word matching tasks. The arrows in Figure 3-2 identify the conditional relations tested during stimulus selection.

In the oral naming task, an English word appeared on the computer screen and the subject was given the following instructions orally in Spanish and English:

> When you see a word on the screen, I want you to tell me what it is, if you know it. If you don't know it, just say "I don't know".

Word presentation was controlled by the experimenter as described in the section on stimulus presentation. Words that were identified by the subject, see word-say word correctly, were eliminated. However, if the subject "sounded out" a word, he was asked to say the word in Spanish. If he could not, the word was included in the next step of stimulus selection as were words not identified by the subject.

Before the first picture/word matching trial, the subject was taught to use the mouse (see Appendix C). Once the subject could reliably move the cursor into all windows,

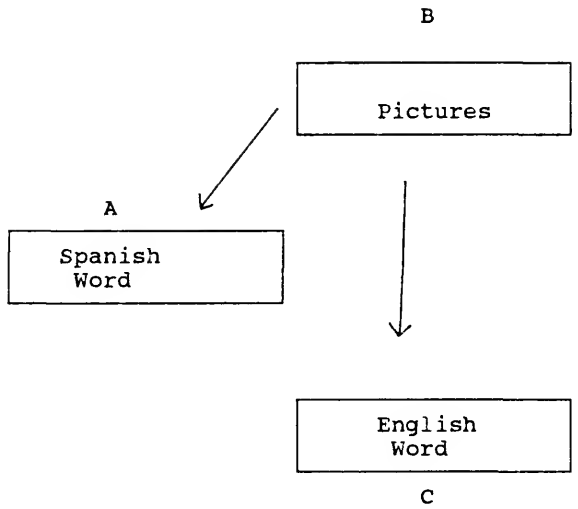


Figure 3-2. Stimulus Sets used in Experiment 1.

Note. Arrows point from the sample stimulus to the comparison stimulus used in the picture/word matching tasks during stimulus selection.

the testing began. The subject received the following oral instructions in Spanish and English:

> You will see a picture in the center of the screen. Look at the picture, put the hand on the picture and push the mouse button. Words will appear. Put the hand on the word that goes with the picture. If you do not know which one is correct, take a guess.

On the first trial, the subject was given feedback ("That's correct") for placement of the hand on the sample, clicking the mouse button, and for clicking on a choice window (without regard to whether a correct choice had been made). No feedback was given in subsequent trials.

In the first matching task, the subject was presented with a picture as the sample and was asked to select the English word that identified the picture from three comparison choices. Words correctly matched were eliminated. Words incorrectly matched were included in the second matching task.

In the second matching task, the subject was presented with a picture as the sample and was asked to select the Spanish word that identified the picture from three comparison choices. Words incorrectly matched were eliminated.

Sessions during this phase were separated from each other by no less than 30 minutes and continued until 20 words had been identified. At the end of each session, the subject was praised for his participation. During the

stimulus selection phase, all correct and error responses were consequated as described for test trials.

Baseline. Figure 3-3 shows the six conditional relations specifically tested during the baseline condition. These tests consisted of: The picture presented as a sample stimulus and the English words presented as comparisons (B-C); the English word presented as a sample stimulus and pictures presented as comparisons (C-B); the Spanish word presented as a sample stimulus and pictures presented as comparisons (A-B); the picture presented as a sample stimulus and Spanish words presented as comparisons (B-A); the Spanish word presented as a sample stimulus and English words presented as comparisons (A-C); and the English word presented as a sample stimulus and Spanish words presented as comparisons (C-A). Each stimulus word was presented as a sample once in each matching task.

During the baseline phase, tests were sequenced so that tests defining the equivalence relation were presented first. That is, the tests of the C-B and B-C relations were presented first. This order was chosen in an attempt to minimize any possibility of learning the equivalence relation from the other tests. Previous references (Haywood et al., 1990; Stoddard & McIlvane, 1985) suggest that sequencing tests from simple to complex in post assessment (i.e., known relations are presented first and the

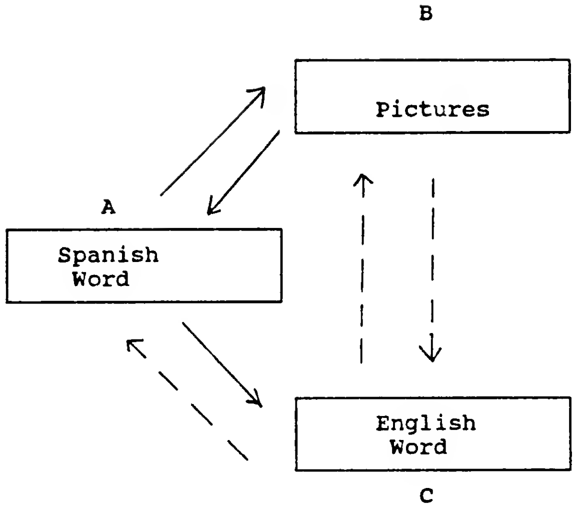


Figure 3-3. Schematic representation of Relations tested during the Baseline Phase in Experiment 1.

Note. Solid arrows denote relations known or taught during the experimental phase. Broken arrows denote equivalence relations only tested.

equivalence test is presented last) enhances the emergence of the equivalence class.

One matching task was presented in each baseline session until all six tests had been completed. One session was scheduled for each class period and sessions were separated by no less than 30 minutes. No training was implemented for two weeks. Then all visual matching tasks for the 20 stimulus words were again assessed.

No instructions were given during the baseline phase and consequences were as described for test trials. At the end of each session, the subject was praised for his participation.

Training procedure

The 20 stimulus words were randomly divided into two lists of 10 words each. (See Table 3-1.) One list was randomly selected to be taught during the first teaching phase. During the second teaching phase, the second list was added and all twenty words were taught.

During each training session, English words from the list were presented as samples and Spanish words were presented as comparisons (match-to-sample task C-A in Figure 3-3). Training sessions for the first 10 words consisted of two presentations of each word in the list (20 total stimulus presentations). The second set of training

Table 3-1

English and Spanish words used in Experiment 1

List 1a	List 1b
cab - taxi	potato - papa
honey - miel	clothing - ropa
jacket - chaqueta	can - lata
dessert - postre	saw - sierra
hen - gallina	knife - cucharo
storekeeper - tendero	sink - lavado
heart - corozon	vase - florero
sausage - salchicha	doll - muneca
mirror - espejo	skillet - sarten
furniture - muebles	mouth - boca

sessions consisted of one presentation of each word in the list (20 total stimulus presentations).

The order in which the samples were presented was randomly assigned for each session. The subject was considered to have learned the words if he made no more than one error in each of two consecutive sessions. No instructions were given during the training procedure and consequences were as described for training trials. At the end of each session, the subject was praised for his participation.

Mid- and postassessment

Midassessments were given one day after the subject had met criteria on the list of 10 words. On the day that the midassessment was administered, the subject was first

required to perform the matching task presented during training sessions (C-A in Figure 3-3) with no more than 1 error. The subject was then tested for the emergence of equivalence relations of all 20 words in a sequence designed to enhance the emergence of the equivalence class. First, the Spanish word was presented as a sample stimulus and English words were presented as comparisons (A-C, probe for symmetry of the relation that had been taught). Second, the picture was presented as a sample stimulus and English words were presented as comparisons (B-C). Finally, the English word was presented as a sample stimulus and pictures were presented as comparisons (C-B). The last two tests together comprise the stimulus equivalence probes. An equivalence class was deemed to have been formed if there was no more than one error on each equivalence task.

Postassessments were administered one day after the subject had met criteria on the list of 20 words. Testing during the post assessment followed the same sequence described above for mid-assessments.

Each test battery, consisting of visual matches as described above, was completed in one day. Each stimulus relation was presented in one session. Sessions were separated from each other by no less than 30 minutes. No instructions or feedback were given during testing and consequences were as described for test trials. At the end

of each session, subjects were praised for their participation.

Experiment 2

Subjects

Two subjects participated in this study. Subject 1 had participated in Experiment 1. Subject 2 had a history of training with conditional relations and the computer program used in this experiment.

Subject 2 was a male from Puerto Rico who was 11.0 years old when the experiment began. He had been enrolled in a middle school program for bilingual students in Alachua County for four months at the time of this experiment. He spent three class periods a day in the bilingual classroom. Although the subject knew some English vocabulary, his primary method of communication was Spanish.

Upon his admission to the middle school program, school personnel administered two English Language tests, the Dade County Aural Language Survey and the Arlington Oral Language Test. He had not passed any item on these two tests. This subject had attained a FSIQ of 108 on a WISC that was administered in Puerto Rico in 1987. He was diagnosed with mixed dominance, leading to directional confusion, and as having difficulty with spatial relations. Records from Puerto Rico show that he had been working below his level since first grade, in spite of receiving after school

tutoring. However, the only special service he was receiving in the United States was bilingual training.

Apparatus and Setting

Sessions were conducted 5 days per week during school hours in a quiet area of the classroom. Each subject was seated at a table facing a Macintosh SE computer equipped with a hard disk and mouse. The experimenter was positioned behind and slightly to one side of the subject during each session. During testing, three sessions per day were scheduled; one during each class period. During training, sessions were scheduled one or two times per day as described in the result section.

The computer program described in Experiment 1 was modified slightly to permit random display of four of the six comparison windows during each experimental trial. Otherwise, stimulus display, cursor movement, and data collection were as described above.

Experimental Procedure

Stimulus presentation and consequences for correct and error responses were as described in Experiment 1 with one exception. The distractor comparison stimuli were randomly assigned to a sample at the beginning of the experiment. Thereafter, the same distractors always accompanied that sample. The position of the distractor stimuli was randomly programmed to occur.

Stimulus selection

Before the experiment began, twenty words were selected from Words for Students of English - Volume 6 (Rogerson, Esarey, Jasnow, Hershelman, Schmandt, Smith & Snellings, 1988). The list contained five nouns, five adjectives, five adverbs, and five verbs. The adverbs were identified as frequently used by McDougal, Littell English (Glatthorn & Rosen, 1988). A list of synonyms was developed from Webster's New World Dictionary (1983).

All words to be used as stimuli were printed on 3x5 cards and presented to the subjects. They were asked to say the word and identify it by its part of speech. If the word was "noun", "adjective", "adverb", or "verb", they were asked to say what the word meant or give an example. Three words were eliminated because they were correctly identified by the subjects. Both subjects recognized the words "verb" and "adjective", which have the same root in English and Spanish. Both identified a verb as an action word and gave an example. Neither were able to identify the other parts of speech or give an example.

Ten words and synonyms were randomly selected from the list of 17 words remaining with at least two words representing each part of speech. These words comprised list 1. The remaining seven words and synonyms were placed

into list 2. The words used during stimulus selection are contained in Table 3-2.

Baseline

Figure 3-4 shows the three stimulus sets and the six conditional relations tested during the baseline condition. All matching tests were assessed for list 1. Once this was completed, all matching tasks were assessed for list 2. Tasks were assessed in the following order: A sample word from set B to be matched to a comparison from set C; a sample word from set C to be matched to a comparison from set B; a sample word from set A to be matched to a comparison from set B; a sample word from set B to be matched to a comparison from set A; a sample word from set A to be matched to a comparison from set C; a sample word from set C to be matched to a comparison from set A.

Baseline sessions consisted of three presentations of each word in the list. There were 30 total presentations for list 1 and 21 total presentations for list 2. The order in which samples were presented was randomly assigned except that each presentation of the same sample was separated from any other presentation of that sample by at least two other words from the list.

One matching task was presented in each baseline session until all six tests had been completed. One session was scheduled for each class period and each session was

Table 3-2

Stimulus words used in Experiment 2 - List 1

Set A	Set B	Set C
quiet	peaceful	adjective
smart	intelligent	adjective
work	toil	verb
often	frequently	adverb
smell	sniff	verb
bravery	courage	noun
bad	harmful	adjective
quite	completely	adverb
chore	task	noun
burden	duty	noun

List 2

Set A	Set B	Set C
empty	unoccupied	adjective
project	proposal	noun
rather	somewhat	adverb
cling	adhere	verb
so	then	adverb
brilliant	vivid	adjective
grow	raise	verb

Words eliminated as known

Set A	Set B	Set C
initiate	begin	verb
reporter	journalist	noun
nearly	almost	adverb

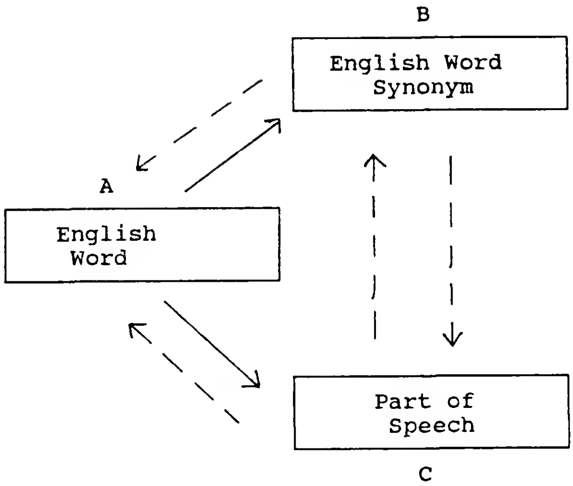


Figure 3-4. Schematic representation of Relations tested during the Baseline Phase in Experiment 2.

Note. Solid arrows denote relations known or taught during the experimental phase. Broken arrows denote equivalence relations only tested.

separated from any other session by no less than 30 minutes. No training was implemented for two weeks. Then all visual matching tasks were again assessed for List 1.

No instructions were given during the baseline phase. Consequences were programmed by the computer as described for test trials. In addition, subjects' responses were reinforced approximately every fifth presentation (Variable Ratio 5 schedule) for placement of the hand on the sample and clicking the mouse button. At the end of each session, subjects were praised for their participation.

Training procedure

The 10 stimulus words from set A were used as samples and stimulus words from set B were used as comparisons during the first training phase. In the second phase, the 10 stimulus words from set A were used as samples and stimulus words from set C were used as comparisons. Training procedures were as described in Experiment 1.

Training sessions in both phases consisted of two presentations of each word in the sample list (20 total stimulus presentations). The order in which the samples were presented was randomly assigned for each session. The subjects were considered to have learned the words if they made no more than one error in two consecutive sessions. No instructions were given during the training procedure and consequences were as described for training trials. At the

end of each session, the subject was praised for his participation.

Postassessment

The basic procedures used during this phase were the same for both subjects. Differences in sequence or additional events will be described in the results section.

Post tests were given one day after the subjects had met criteria on the trained relationship. On the day that post tests were given, the subjects were first required to perform the matching tasks presented during training sessions (A-B and A-C relations in Figure 3-4) with no more than 1 error. The subjects were then tested for the emergence of equivalence relations in the following sequence. First, a sample word from set B to be matched to a comparison from set A; a sample word from set C to be matched to a comparison from set A; a sample word from set B to be matched to a comparison from set C; a sample word from set C to be matched to a comparison from set B. The B-C and C-B tests comprise the stimulus equivalence probes. An equivalence class was deemed to have been formed if the subject correctly matched a sample and comparison stimulus for 3 of 3 stimulus presentations in both the B-C and C-B test.

During post assessment, each matching task was presented in one session. Sessions were separated from each other by no less than 30 minutes. No instructions or feedback were given during post assessment and consequences were as described for test trials. At the end of each session, subjects were praised for their participation.

Maintenance

Two weeks and four weeks after the last post test, the subjects were retested on all matching tasks. All post test procedures were followed. No instructions were given during this testing and consequences were as described for test trials. At the end of each session, subjects were praised for their participation.

Treatment of the Data

The percent of correct choices was the primary datum upon which conclusions about the presence or absence of equivalence relations were based. At the end of each test session, the percent of correct responses was computed for each matching task by dividing the number of correct responses by the total number of trials. These data were then displayed graphically by phase. Graphic display permitted descriptive comparisons of correct responses across pretest, post test and maintenance phases.

During the training phase, the percent of correct responses was computed at the end of each trial block and was compared to the criterion for moving to the next sequence of words. This percent was recorded for later use in calculating the number of training trials required for each subject to learn all words. These data were also displayed graphically to permit comparisons across training phases.

In experiment 2, a ratio was also computed consisting of the number of trials in which each word was correctly matched over the number of possible trials. This ratio provided additional data relative to the formation of an equivalence class.

Reliability of test data was assessed during Experiment 1 only. The experimenter recorded correct and error responses during 20% of the sessions in each phase. Because there was 100% agreement between the experimenter-recorded data and the computer-recorded data, it was not deemed necessary to continue reliability checks.

CHAPTER IV
RESULTS

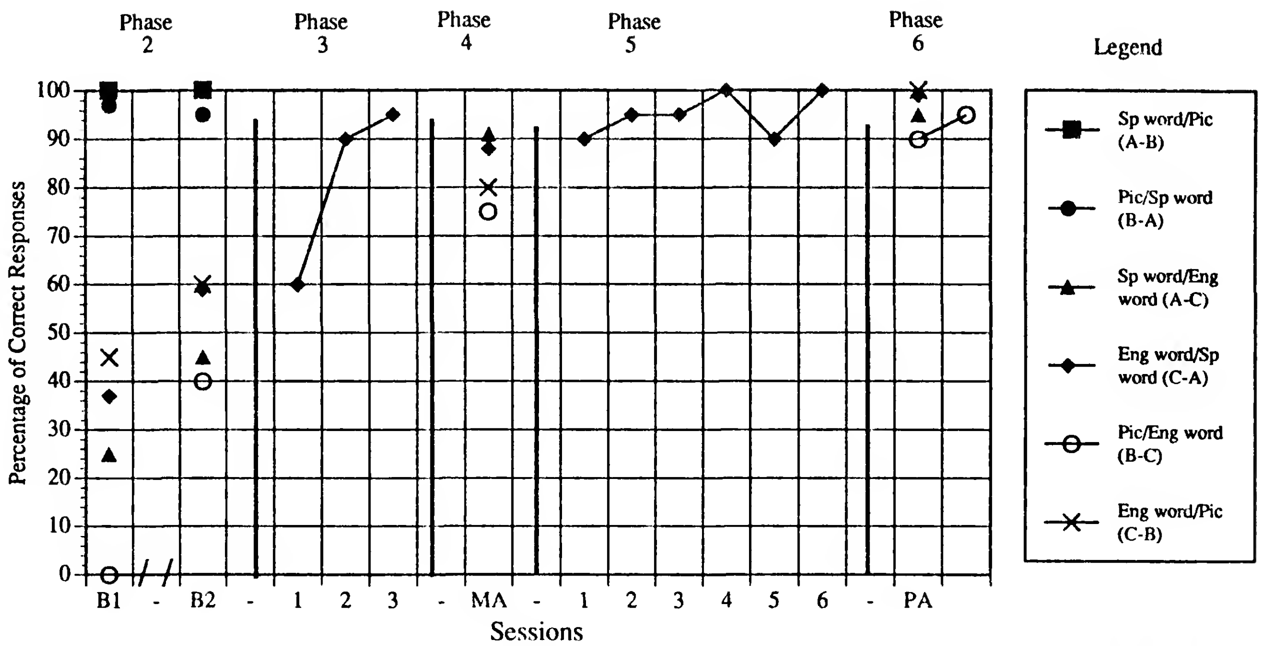
Experiment 1

Experiment 1 was designed to answer the first question: Will a limited English speaking subject, who is able to match Spanish words to pictures, demonstrate the ability to match English words to pictures after he is taught to match Spanish words to English words?

Baseline

Baseline test results are presented in Phase 2 of Figure 4-1. During the first battery of baseline tests, the subject was able to correctly match all of the stimuli presented during tests of the A-B relationship in which a Spanish word was presented as a sample and pictures were presented as comparisons, and tests of the B-A relationship in which a picture was presented as a sample and Spanish words were presented as comparisons. One hundred percent correct responding to stimuli presented in tests of the B-A relationship was one criterion for stimulus selection.

The subject was able to correctly match only 37% of the words during the test of the relationship to be taught, matching Spanish word comparisons to an English word sample



Note. // denotes 2 weeks between sessions
B dentoes Baseline
MA and PA denote Mid- and Postassessment

Figure 4-1. Experiment 1: Percentage of correct responses across Phases 2 through 6. All stimuli were tested in Phases 2,4, and 6. Ten words were trained in Phase 3. All 20 words were trained in Phase 5.

(C-A). The symmetrical relation of this task, matching English word comparisons to a Spanish word sample was even lower at 25% correct matches.

No correct matches were recorded in the test of the B-C relation, matching English word comparisons to a picture sample. This score was a criterion for stimulus selection. However, the score of the test on the symmetrical relation, matching picture comparisons to an English word sample, was high at 45% correct matches. The scores on these two tests provide the critical data for assessing the emergence of stimulus equivalence.

During the second battery of baseline tests, scores on all but one of the matching tests increased. The decrease in correct matches during the test of the B-A relationship, matching Spanish word comparisons to a picture sample, from 100% correct to 95% correct represents one incorrect match. The subject achieved correct matches for 60% of the words in the relationship to be taught, matching Spanish word comparisons to an English word sample. The symmetrical relation of this task, matching English word comparisons to a Spanish word sample was lower at 45% correct matches.

Data for the B-C and C-B relations show increases to 40% and 60% correct respectively. Of these correct responses, 2 stimulus pairs from list 1 and 5 stimulus pairs

from list 2 were correctly matched in both the B-C and C-B tests.

If all stimuli are unknown, there is a 33% chance probability of correctly matching a sample to a comparison stimulus when three possibilities are presented. If one or two stimuli are known, the probability increases to 50% and 100% respectively. During the second baseline phase, all test scores exceeded the 33% chance level. In addition, a total of seven of the 20 word pairs that were presented were correctly matched in the B-C and C-B tests.

These data suggest that the subject was not responding at chance levels but had formed relationships among at least seven stimulus sets outside of the experimental session. It is unknown whether these data represent the formation of equivalence classes, as might be indicated by increased percent correct responding on the A-C and C-A tests, or direct training of the B-C and C-B relationship outside of the experimental session. If directly trained, the assumption of equivalence may not be made.

Training and Assessment Data

Three sessions were required for the subject to meet criterion on list 1a. At the end of training the subject was responding at 95% accuracy on list 1a.

Scores for the four tested relations increased over baseline data during the mid assessment. The subject achieved a score of 90% correct on both the taught relation, C-A, and the symmetrical relation, A-C. Scores on the two equivalence tests increased to 75% for the B-C relation and 80% for the C-B relation. Of these correct responses, 7 stimulus pairs from list 1 and 7 stimulus pairs from list 2 were correctly identified in both the B-C and C-B tests. These tests indicate that the subject had formed a relationship among seven additional stimulus sets as compared to baseline responding.

Six additional sessions of training were required for the subject to master the criterion for all 20 words. The subject began training with 90% correct responding. At the end of training the subject was responding at 100% accuracy.

The subject achieved a post assessment score of 100% correct for the taught relation, C-A, and for the transitive relation, C-B. A score of 95% was achieved for the symmetrical relation A-C. Initially, the subject only achieved a score of 90% correct on the equivalence test, B-C. However, when a second test was administered at the end of the post assessment phase, the score increased to 95% correct.

Although this subject achieved post assessment scores of 95% correct on both tests that were designed to assess the formation of equivalence classes, it was not clear that equivalence was the basis for these scores. Increases in percent correct overall responding and in stimulus pairs correctly matched across both the B-C and C-B tests could have been due to extra-experimental contact with the stimuli. Experiment 2 was designed to reduce the probability of contact with the experimental stimuli outside of the experimental session, while continuing to assess the acquisition of cognitively complex academic skills without resorting to the use of arbitrary stimuli.

Experiment 2

Experiment 2 was designed to answer two questions related to the formation of equivalence classes within cognitively complex subject matter, English grammar. First would subjects form an equivalence relation among words and their part of speech. Second, would the relations maintain after two and four weeks?

Experimental sessions varied between the subjects due to the rate at which they met the training criterion and differences in equivalence class formation. Results for each subject are presented separately.

Subject 1

Baseline

The sequence of tests is presented in Appendix D. Each test was administered in one session and no test was repeated during this phase.

Baseline test results are presented in Phase 1 of Figure 4-2. During the first battery of baseline tests, the subject scored poorly on the four relations that included the part of speech as one member of the set. A score of 27% correct was obtained on both the A-C relation and its symmetrical counterpart C-A. Scores were even lower on the two equivalence tests, B-C and C-B. The subject obtained scores of 17% and 23% respectively. The scores on these two tests (B-C and C-B) provide the critical data for assessing the emergence of stimulus classes.

The percent of correct responses was higher on the two relations that required the subject to match a word and its synonym, A-B and B-A. The subject obtained a score of 33% correct on the A-B relation and a score of 50% correct on the B-A relation.

Previous stimulus equivalence experiments have not identified an absolute level of responding at which a stimulus relation is considered to have been formed. However, the criterion is usually very conservative, and ranges from 90% to 100% correct responding. In this

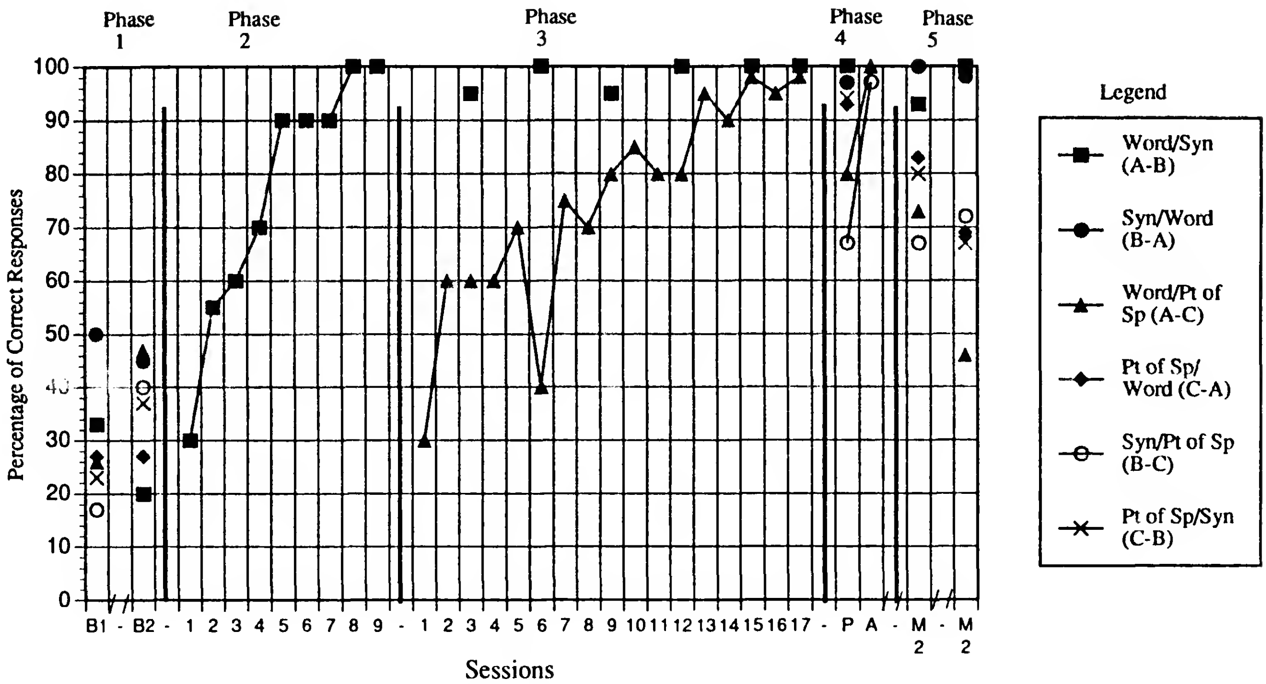


Figure 4-2. Subject 1. Percentage of Correct Responses across Phases 1 through 5. Phases 1, 4, and 5 consisted of testing. Conditional relations were trained in Phases 2 and 3.

Note. // signifies two weeks between data collection.
B denotes Baseline
PA denotes Postassessment
M denotes Maintenance

experiment, the subject was considered to have formed a relation between stimuli in a set if he correctly matched the sample and comparison stimuli in all test presentations (3).

Table 4-1 compares the overall percent of correct responding within each relation to the number of words correctly matched in 3 of the 3 stimulus presentations. During the first baseline tests, the subject only matched one stimulus pair correctly three of three presentations. This occurred in the B-A test in which the subject achieved his highest overall percent correct score. These data provide additional support for the conclusion that no relation had been established in spite of high levels of correct responding (50% in the B-A relation).

Figure 4-2 shows that during the second battery of baseline tests variability across relations was reduced from the first baseline tests. Median percent correct responding across relations increased from 27% to 37%. Overall percent correct scores increased for three of the tested relations. The subject obtained a score of 40% correct on the B-C relation, 37% correct on the C-B relation, and 46% correct on the A-C relation.

The subject also increased the number of stimulus pairs correctly matched in these three relations from 0 to 1. The same word pair, "chore-noun", was correctly matched

Table 4-1

Comparison of percent correct scores and number of word pairs correctly matched on 3 of 3 stimulus presentations within each relation - Subject 1

Phase		Baseline 1	Baseline 2	Post Assessment	Maintenance 2 week	Maintenance 4 week
A-B	%	33	20	100	93	100
	#	0	0	10	9	10
B-A	%	50	46	97	100	100
	#	1	1	9	10	10
A-C	%	27	46	80 97	73	46
	#	0	1	8 9	6	1
C-A	%	27	27	93	83	70
	#	0	0	9	7	5
B-C	%	17	40	67 97	67	70
	#	0	1	7 9	6	6
C-B	%	23	37	93	80	67
	#	0	1	9	7	4

Note. % denotes percent of correct responding
# denotes number of word pairs correctly matched on 3 of 3 stimulus presentations within a test

in tests of both the B-C and C-B relation. Chore, noun, and task were replaced by a set of words randomly selected from list 2, unoccupied, adjective and empty. (See Table 3-2).

A slight decrease, from 50% to 46% correct, was noted in the score of the B-A test. The subject again correctly matched one word pair in all three of the three stimulus presentations during this test. However, the word pair correctly matched differed from the one he had correctly matched in the first baseline test.

The score on the test of the A-B relation decreased to 20% correct, while the score of the C-A relation remained the same at 27%. No word pairs were correctly matched on all three stimulus presentations during tests of these two relations.

Training

The sequence of training is presented in Appendix D. Nine sessions were required for this subject to meet criterion on the A-B relation. Data are presented in Phase 2 of Figure 4-2. At the end of training, he had achieved 100% correct responding. Accuracy of responding was maintained during the training phase of the A-C relation as noted in Figure 4-2.

Training proceeded at a slower rate for the A-C relation, 3.5% per day as compared to 7% per day improvement (experimental phase 3). There was slightly more variability

in responding across sessions, and seventeen sessions were required to meet criterion. At the end of training, he had again achieved 100% correct responding.

Review sessions of both the A-B and the A-C relations were conducted one day after criterion had been met on the A-C relation. A score of 100% correct was obtained on both.

Postassessment

Phase 4 data in Figure 4-2 show that scores increased over baseline levels for all relations tested during the post assessment. The subject achieved a score of 100% correct on the test of the A-B relation and 97% correct on the test of its symmetrical counterpart, the B-A relation. This score of 97% represents one error overall. Ten word pairs in the A-B relation and 9 word pairs in the B-A relation were correctly matched on three of three stimulus presentations.

A score of only 80% correct was achieved on the initial test of the A-C relation. Eight word pairs were correctly matched on three of three stimulus presentations. Consequently, one training review session was presented. Following this review, the A-C test was repeated in a separate session. On this second test, the subject responded correctly 97% of the time. Nine word pairs were correctly matched on three of three stimulus presentations.

The subject missed one presentation of the sample word "smell".

The C-A relation was tested following the second test of the A-C relation. The subject achieved a score of 93% correct. Nine word pairs were correctly matched on three of three stimulus presentations. The subject missed two presentations of the sample word "adverb" when "quite" was the correct comparison choice.

A score of only 67% correct was achieved upon the initial test of the B-C relation. Seven word pairs were correctly matched on three of three stimulus presentations. This test was followed by the C-B test on which the subject achieved a score of 93% correct. The score of the C-B test represents nine word pairs correctly matched on three of three stimulus presentations. Without further training or review, the B-C test was repeated. The subject achieved a score of 97% correct, and correctly matched nine word pairs on three of three stimulus presentations.

Although nine word pairs were correctly matched in each of the B-C and C-B tests, only 8 word pairs were correctly matched across both tests. As per the criterion noted above, the eight word pairs correctly matched across the two tests are considered to represent selections that are consistent with the expansion of classes and the formation of an equivalence relationship.

<u>Maintenance</u>

Maintenance test results are presented in Phase 5 of Figure 4-2. The first battery of maintenance tests was conducted two weeks following the post assessment. Scores decreased on five of the six tested relations. In addition, scores for the symmetrical relationships were higher than those achieved in the relations that had been taught.

The subjects score on the test of the A-B relation only decreased slightly to 93% correct. Nine word pairs were correctly matched on three of three stimulus presentations. A slight increase, to 100%, was noted in the test of the B-A relation.

The score of the A-C relation decreased to 73%. In addition, word pairs correctly matched on three of three stimulus presentations decreased from a high of nine, achieved during post assessment, to six during this maintenance test. A score of 83% was achieved on the test of the C-A relation. Seven word pairs were correctly matched on three of three stimulus presentations. The scores on the tests of the equivalence relations also decreased to 67% correct on the B-C relation and 80% correct on the C-B relation. These scores represent six and seven word pairs, respectively, that were correctly matched on three of three stimulus presentations.

There was some consistency in the words missed across the six tests. Word pairs from the stimulus set "bravery-courage-noun" and "often-frequently-adverb" accounted for 62% of the incorrect matches in tests of the A-C, C-A, B-C,and C-B relations.

Some further decreases in scores are noted when tests were again administered four weeks following the post assessment. Figure 4-2 shows that three test scores are lower at the four week maintenance check than they were at the 2 week maintenance check. Scores on the symmetrical relations are the same or higher than those recorded for the relations that were taught. In addition, the lowest score is recorded for the A-C relation.

The highest scores were again achieved on the tests of the A-B and B-A relationship. The subject correctly matched 100% of the word pairs on both tests.

The subject achieved a score of only 46% correct on the A-C relation, a relation that had been trained. Only one word pair was correctly matched in three of three stimulus presentations. This score is comparable to baseline levels of responding. In contrast, a score of 70% correct was achieved on the C-A relation. Five word pairs were correctly matched in three of three stimulus presentations. The word pair correctly matched in the A-C relation,

"completely-quite", was consistently missed during the test of the C-A relation.

The score for the test of the B-C relation maintained at 67% correct and six word pairs correctly matched in three of three presentations. The score of the C-B relation decreased to 67% correct. However, only four word pairs were correctly matched in three of three presentations. This represents a loss of 3 word pairs from the 2 week maintenance follow up.

Consistency in word pairs missed was noted among tests of the relations and between maintenance tests. During this maintenance check, word pairs from the stimulus set "bravery-courage-noun" and "often-frequently-adverb" were again incorrectly matched in tests of the A-C, C-A, B-C, and C-B relations. These words were never correctly matched in tests of the A-C and B-C relations. Word pairs from the stimulus set "quite-completely-adverb" were incorrectly matched in tests of the C-A, B-C, and C-B relations; although this stimulus set accounted for the only word pair correctly matched three of three presentations during the A-C relation.

The subject also displayed consistent responding between tests of the C-A and C-B relations. Four of the word pairs correctly matched on three of three stimulus presentations were the same in both tests.

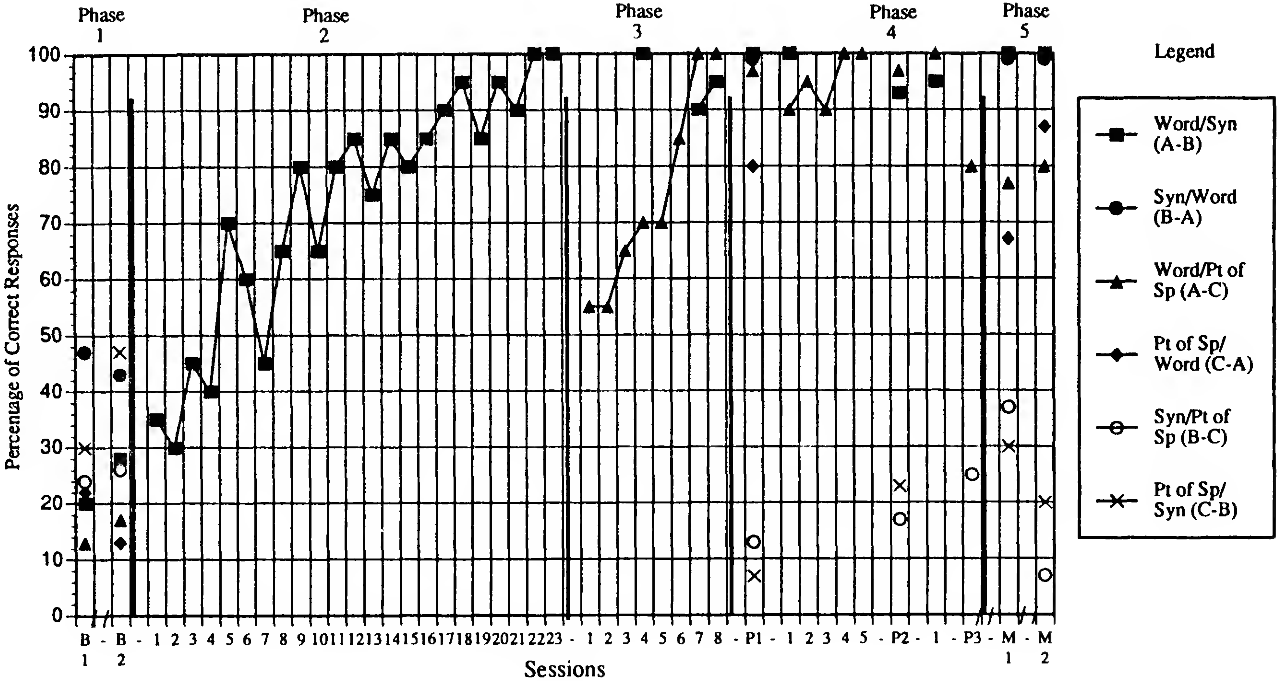
Subject 2

Baseline

The sequence of relationships tested and trained within each experimental phase is presented in Appendix D. During the baseline phase, each test was administered in one session and no test was repeated.

Baseline test results are presented in phase 1 of Figure 4-3. Scores reflect overall incorrect and inconsistent responding on five of the six baseline tests in the first battery. The subject responded correctly during 13% of the presentations on the test of the A-C relation and 23% of the presentations on the test of its symmetrical counterpart, C-A. Scores on the two equivalence tests, B-C and C-B, were also low, at 23% and 30% respectively. Similarly, a score of only 20% correct was obtained on the test of the A-B relation.

Table 4-2 provides additional data related to inconsistent responding across the first baseline tests. Only one stimulus pair was correctly matched in three of three stimulus presentations. This occurred during the test of the B-C relation. Although overall correct responding was relatively high during the test of the B-A relation, 47% correct, no stimulus pair was correctly matched in 3 of 3 stimulus presentations.



Note. // signifies two weeks between data collection.
B denotes Baseline
PA denotes Postassessment
M denotes Maintenance

Figure 4-3. Subject 2. Percentage of Correct Responses across Phases 1 through 5. Phases 1, 4, and 5 consisted of testing. Conditional relations were trained in Phases 2 and 3.

Table 4-2

Comparison of percent correct scores and number of word pairs correctly matched on 3 of 3 stimulus presentations within each relation - Subject 2

Phase		Baseline 1	Baseline 2	Post Assessment 1	Post Assessment 2	Post Assessment 3	Maintenance 2 week	Maintenance 4 week
A-B	%	20	27	100	93		100	100
	#	0	1	10	9		10	10
B-A	%	47	43	100			100	100
	#	1	2	10			10	10
A-C	%	13	17	100	97	80	77	80
	#	0	0	10	9	8	6	8
C-A	%	23	13	80			67	87
	#	0	0	7			5	8
B-C	%	23	27	13	17	25	37	7
	#	1	0	0	0	0	2	0
C-B	%	30	47	7	23		30	20
	#	0	1	0	2		1	1

Note. % denotes percent of correct responding
# denotes number of word pairs correctly matched on 3 of 3 stimulus presentations within a test

When baseline tests were repeated, increased correct responding was noted in tests of four of the six relations; A-B, A-C, B-C, and C-B. The subject obtained a score of 27% correct on the test of the A-B relation, 17% correct on the test of the A-C relation, 27% correct on the test of the B-C relation, and 47% correct on the test of the C-B relation. Scores on the tests of the remaining two relations, B-A and C-A, slightly decreased to 43% and 13% respectively.

Phase 1 of Figure 4-3 shows that changes in percent of correct responding, whether increased or decreased, represent rather minor differences on five of the tests. That is, changes reflect differences of only one or two trials. The overall increase resulted in only a slight increase in the median percent of correct responding from 23% to 27%. In addition, Table 4-2 shows that stimulus pairs correctly matched in 3 of 3 presentations increased in only three of the tests, A-B, B-A, and C-B. No stimulus pair was consistently matched across tests.

The score obtained on the test of the C-B relation presents a possible problem in that it reflects an increase in correct responding on seven trials as well as an increase in the number of stimulus pairs correctly matched from 0 to 1. If this word pair, "adjective-harmful", had also been correctly matched in the B-C test, one might suspect that a conditional relation had been formed and the word pair would

have been replaced. However, the word pair was not correctly matched in any other test and was, therefore, not replaced.

Training

The sequence of training was the same as that provided for Subject 1. First the A-B and then the A-C relation was trained. However, during training of the A-B relation, sessions were increased to two per day beginning on the fifth day. Twenty three sessions, occurring over 14 days, were required for this subject to meet criterion on the A-B relation. At the end of training, he had achieved 100% correct responding. Correct responding was maintained between 90% and 100% during the A-C training phase as noted in Figure 4-3.

Training sessions were decreased to one per day and learning proceeded at a much more rapid rate for the A-C relation, 10% per day as compared to 4% per day improvement. Only eight sessions were required for this subject to meet criterion. At the end of training, he had again achieved 100% correct responding.

A review session of the A-B relation was conducted the day that criterion had been met on the A-C relation. A score of 95% correct was obtained. This score meets the criterion previously set for post assessment to occur.

Postassessment

Phase 3 of Figure 4-3 presents post assessment scores for this subject. The first post assessment consisted of tests of all six relations begun one day after the subject had met criterion on the A-C relation. Four test scores were increased over baseline. The subject achieved a score of 100% correct on the tests of the two trained relations, A-B and A-C, and the test of the symmetrical relation, B-A. However, he only responded correctly 80% of the time on the symmetrical relation, C-A. Seven word pairs in the C-A relation were correctly matched on three of three stimulus presentations. Based upon previous research, the formation of an equivalence class might be expected for only the seven sets of stimuli for which symmetry had been demonstrated, i.e. those that were correctly matched across the four tests.

The percent of correct responding decreased in tests of both the B-C and C-B relations when compared to baseline levels. A score of 13% correct was achieved on the test of the B-C relation and a score of 7% correct was achieved on the test of the C-B relation. No word pairs were correctly matched on three of three stimulus presentations during tests of either relation. In spite of apparent symmetrical responding on sever word pairs, no equivalence class formation was demonstrated.

Although this subject's responses were incorrect, they were not random. When presented with a sample stimulus, he was likely to chose the same comparison on all three trials. Over the two tests, B-C and C-B, consistent matching of sample and comparison stimuli occurred on 72% of trials. A review of responses on each trial showed that responses were proportionally distributed across the parts of speech in the B-C test and across words in the C-B test.

In previous studies, failure to form equivalence classes has been remediated by retesting and retraining if necessary (Bush et al, 1989; McDonagh et al., 1984; Sidman & Tailby, 1982). Consequently, training trials for both the A-B and A-C relations were implemented. These trials were exactly the same as those described above in the Training Phase.

One training review session was presented for the A-B relation. The subject responded correctly on 100% of the trials. On the first review of the A-C relation, the subject only responded correctly on 90% of the trials. Four additional sessions were required for the subject to reach criterion. Following this review, the second series of post assessment tests was initiated. In this series only four relations were tested, A-B, A-C, B-C, and C-B.

The subject responded correctly on 93% of the trials during the test of the A-B relation and on 97% of the trials during the test of the A-C relation. Nine word pairs were correctly matched on three of three stimulus presentations.

Responses on the equivalence tests were only somewhat improved over the first post assessment and remained slightly lower than baseline levels of responding. The subject responded correctly on 17% of the trials during the test of the B-C relation and on 23% of the trials during the test of the C-B relation. Although two word pairs were correctly matched on three of three stimulus presentations in the test of the C-B relation, no word pairs were correctly matched on 3 of 3 stimulus presentations in the test of the B-C relation. Correct C-B matches may have been due to the subject's practice of consistently choosing the same comparison rather than the formation of equivalence classes.

Consistent matching of sample and comparison stimuli again occurred on 66% of trials during tests of the B-C and C-B relations. However, during this post assessment, the subject was more likely to choose certain comparison stimuli across all trials. That is, the subject chose "adjective" as the comparison in 50% of the trials when presented with a sample stimulus during the B-C test; he chose "completely"

as the comparison in 37% of the trials when presented with a sample stimulus during the C-B test.

In a further effort to increase the opportunity for the formation of equivalence classes, a test was designed in which B-C test trials were mixed with A-C test trials. Sidman et al. (1982) suggested that mixing test and training trials might enhance the emergence of symmetry and lead to equivalence class formation. The test consisted of 30 stimulus presentations. Ten stimulus pairs from the A-C test, one presentation of each pair, and ten stimulus pairs from the B-C test, two presentations of each pair, were included.

This test disrupted correct responding to the stimuli from the A-C test. The subject only responded correctly to 80% of the test trials. Correct responding to the stimuli from the B-C test was slightly increased over the previous two post assessments, but remained at baseline levels of responding. No sample and comparison were correctly matched on more than one trial. No further testing was undertaken in this phase for this subject.

Maintenance

Maintenance test results are presented in Phase 5 of Figure 4-3. The first battery of maintenance tests was conducted two weeks following the last post assessment. Scores decreased on two of the six tested relations.

The subject responded correctly 100% of the time on the test of the A-B relation and its symmetrical counterpart, B-A. This represents ten word pairs correctly matched on three of three stimulus presentations.

The score of the A-C relation decreased to 77%. In addition, word pairs correctly matched on three of three stimulus presentations decreased from a high of ten, achieved during the first post assessment, to six during this maintenance test. The score of the C-A test also decreased to 67%. Five word pairs were correctly matched on three of three stimulus presentations.

There was some consistency in the words missed across the A-C and C-A tests. Word pairs from the stimulus set "bravery-courage-noun", "smell-sniff-verb", "bad-harmful-adjective" and "often-frequently-adverb" were responded to incorrectly on tests of both the A-C and C-A relations.

Although the scores on the tests of the equivalence relations increased, the data do not support the conclusion that any equivalence class was formed. The subject responded correctly 37% of the time on the B-C relation and correctly matched two word pairs on three of three stimulus presentations. A score of 30% correct was achieved on the C-B relation and one word pair was correctly matched on three of three stimulus presentations. The word pairs correctly matched differed between the two tests. In

addition, the set consisting of "frequently-often-adverb" was correctly matched in the test of the B-C relation but was responded to incorrectly when presented in tests of the A-C and C-A relations.

The subject continued to respond consistently when presented with a sample stimulus, although he did not always choose the correct comparison. During the tests of the A-C and C-A relations, a sample was matched with the same comparison stimuli in 90% of the trials. During the tests of the B-C and C-B relations, a sample was matched with the same comparison stimuli in 76% of the trials.

Four weeks following the post assessment correct responding increased in tests of the A-C and C-A relations. The subject achieved a score of 80% correct on the test of the A-C relation and 87% correct on the test of the C-A relation. Eight word pairs in each test were correctly matched on three of three stimulus presentations. However, only five word pairs were correctly matched across both tests. One hundred percent correct responding was again recorded for tests of the A-B and B-A relations.

The score for the test of the B-C relation decreased to 7% correct with no word pairs correctly matched in three of three presentations. The score of the C-B relation decreased to 20% correct. One word pair was correctly matched in three of three presentations. These scores are

again below baseline levels. During these tests the subject consistently chose a specific comparison when presented with a sample stimulus on 76% of the trials.

Word pairs responded to incorrectly were not consistent between the A-C and C-A tests. "Quite" and "quiet" were confused during the A-C test. That is, the subject consistently chose "adjective" when "quite" was the sample and "adverb" when "quiet" was the sample. "Bravery" and "smart" were consistently missed during tests of the C-A relation. The only word pair missed at both the two and four week follow-up tests was "bravery-noun".

Summary of Results

In Experiment 1, post assessment scores for all relations were at criterion levels of 95% correct or better. The subject was able to match English words to pictures (C-B), pictures to English words (B-C), Spanish words to English words (A-C), and English words to Spanish words (C-A). However, these high scores only verified the conditional discriminations. Increases in scores during the second battery of baseline tests suggest that some of the conditional relations in the equivalence tests (B-C and C-B) were in the process of being learned. If the conditional relation is taught or reinforced, the assumption of the emergence of equivalence may not be made.

Stimuli used in Experiment 2 were chosen in an effort to eliminate possible contact outside of the experimental setting. It appeared that this tactic was successful in that only one conditional relation might have been formed by one subject during baseline testing. To protect against this possibility, the words in this conditional relation were replaced.

Neither subject responded randomly during baseline testing. That is, subjects consistently chose a particular comparison in the presence of a particular sample. Saunders et al. (1988) have previously identified this phenomena in mentally retarded subjects.

Both subjects learned the two relations that were trained: matching words to their synonyms (A-B) and matching words to their parts of speech (A-C). However, post assessment data were very different for each subject. Subject 1 met criterion on the tests of the equivalence relations: matching the synonyms to their parts of speech (B-C) and matching the parts of speech to synonyms (C-B). Tests of the symmetrical relations were also at criterion levels.

In contrast, Subject 2 did not form equivalence relations in spite of repeated training and testing. Scores on tests of the B-C and C-B relations were at or below baseline levels. This subject's scores on the C-A test for

symmetry, matching the part of speech to a word, were also below criterion. However, seven word pairs were correctly matched on three of three presentations within this test. The subject was responding in a manner consistent with symmetry on at least these seven word pairs.

The two subjects also differed on their maintenance tests. At two and four week follow-up tests, Subject 1 maintained criterion level responding on one trained relation (A-B) and the symmetrical relation (B-A). Two weeks following the post assessment, tests of the equivalence relations (B-C and C-B), the test of the second trained relation (A-C), and the test of the symmetrical relation (C-A) were below criterion levels but higher than baseline responding. These four tests had the part of speech as one member of the word pair. Four weeks following the post assessment, correct responding had deteriorated further on the four relations (B-C, C-B, A-C, and C-A). Responses to the test of the trained relation (A-C), matching the word to the part of speech, decreased to baseline levels. This test score was lower than that of both equivalence tests and the test of the symmetrical relation.

Subject 2 also maintained criterion level responding on the trained relation (A-B) and the symmetrical relation (B-A) at two and four week follow-up tests. Scores on the

tests of the second trained relation (A-C) and the symmetrical relation (C-A) were lower than post assessment scores but remained above baseline levels of responding. Scores were higher during the four week follow-up than they were during the two week follow-up.

Equivalence test scores remained at or near baseline levels, although they were higher at the two week follow-up test than during post assessment or the four week follow-up. The data did not support the conclusion that an equivalence class was formed. However, this subject continued to make consistent conditional selections by choosing the same comparison stimulus when presented with a sample.

CHAPTER V
DISCUSSION

The two experiments described above were designed to investigate whether equivalence classes would be formed within two cognitively complex academic tasks; reading comprehension and English grammar, specifically, parts of speech. A related question examined the retention of the relations, both learned and emergent, in the second task, English grammar.

Since 1971, research has demonstrated that equivalent classes of stimuli may be formed using matching-to-sample procedures. One of the major implications of this research has been its potential for application to a variety of cognitively complex academic subject matter. Investigators have studied linguistic processes, relations under contextual control, sequence responding and expansion of class size. However, they have used arbitrary, experimenter designed stimuli, rather than meaningful stimuli to identify the fundamental features of equivalence. The use of arbitrary stimuli enable the investigator to avoid possible contact of the subject with the stimuli outside of the

experimental setting, spare inevitable failure, and preclude the establishment of error patterns that might later compete for control with relevant instructional stimuli. When academic tasks have been the focus of the research, they have not included complex subject matter such as science, history, or grammar. In addition, only a limited number of studies have addressed maintenance and generalization of emergent relations.

In the classroom, focusing on the formation of equivalence relations may prove to be an efficient and highly productive method of teaching. It is efficient if equivalence classes are formed from purely visual stimuli that may be presented on a computer, thus relieving the teacher for other activities. It is productive if large numbers of untaught relations emerge. However, stimulus equivalence is relatively unknown and unused in the practice of education.

A major impetus for the present study was to examine stimulus equivalence within a task that was directly relevant to teachers but that would also extend current research. The results provide data related to a new subject population, to cognitively complex academic subject matter, and to retention of relations. Incidentally, the study provides for the collection of information related to the use of a computer application in the classroom.

The discussion is organized in sections that correspond to the experimental questions: First, will limited English speaking subjects, who are able to match Spanish words to pictures, demonstrate the ability to match English words to pictures after they are taught to match Spanish words to English words? Second, will limited English speaking subjects form an equivalence class consisting of an English word, its English synonym, and the word that identifies the correct part of speech when they are taught to match a word to its synonym and to match a word to its part of speech? Third, will the performance on the second task be maintained after a two and four week interval? The discussion continues with information related to computer use, followed by research implications, and concludes with suggestions for future research.

Research Findings

Emergence of Reading Comprehension

The subject in Experiment 1 demonstrated the ability to match English words to pictures after he was taught to match Spanish words to English words. However, no conclusions may be drawn relative to the formation of equivalence classes because correct responding increased within the equivalence tests (B-C and C-B) prior to experimental manipulation. As previously explained, it is important that the conditional relation is not reinforced or taught in order to be certain

that equivalence is the basis for the performance. In this experiment, the basis for the performance could not be identified. Therefore, it is not possible to identify increased correct responding as evidence of equivalence classes rather than simply conditional responding based upon some other dimension of the stimuli.

A likely explanation is that the subject learned the conditional relation B-C or C-B. This may have happened in two ways. First, the subject may have encountered the B and C stimuli (English word and picture) in some other setting. Informal feedback from parents at the end of the experiment indicated that the subject had in fact asked what object some words described, e.g., "What is a sink?". A second possibility is that relations were learned as a by-product of the repeated testing procedure. Across both baseline phases, total number of correct matches increased with each subsequent test. K. Saunders (personal communication, November, 1989) noted that subjects may learn to choose a correct comparison when it is presented with distractor stimuli that vary across trials because it is the only stimulus that always occurs with a particular sample. The suggested solution, to always use the same distractors, was implemented in Experiment 2.

These results may be contrasted with other studies in which similar meaningful stimuli were used. Mentally retarded subjects (Sidman & Cresson, 1973) and preschool children (Joyce & Wolking, 1989) demonstrated the emergence of an equivalence class that consisted of words and pictures. Adolescents with traumatic brain injury demonstrated the emergence of an equivalence class that consisted of Spanish words and pictures (Joyce et al., 1990). In two of these experiments, Joyce and Wolking and Joyce et al., only one battery of baseline tests was administered. Therefore, changes prior to training could not be evaluated. However, the subjects in the Sidman and Cresson experiment were tested extensively both before and during training. Moreover, these subjects only responded correctly to the words that had been taught.

The differences noted between this experiment and the ones cited above may be the result of differences in the populations. For example, the extent of the handicaps presented by the mentally retarded subjects and the traumatic brain injury subjects may have precluded learning except in the very structured settings offered during the experiment. Another possibility is that subjects' access to experimental stimuli was limited in the three studies presented above. Preschool children and mentally retarded persons are usually not given experience with printed words.

Neither are Spanish words common in most environments in the United States. A third possibility is that more comparison stimuli are needed. Sidman and Cresson (1973) used five to nine comparisons with each sample presentation. Further study is necessary to determine the effects of stimulus presentation during testing.

Emergence of Grammar Equivalence Relations

Two subjects participated in this experiment. One subject demonstrated the emergence of an equivalence class consisting of a word, its synonym, and the part of speech describing both of them. The second subject did not form an equivalence class although other relations were at, or near, criterion levels.

Other researchers have reported subjects who have failed to demonstrate the emergent relation (Bush et al., 1989, Haywood et al., 1990, McDonagh et al., 1984). However, these subjects are rare. Usually with repeated testing and training, equivalence classes emerge. The failures noted above are different than Experiment 1 in that the number of emergent relations expected have required second order or contextually controlled conditional discriminations or testing has become lengthy and subjects have refused to finish the experiment.

It is uncertain why some subjects fail to form an equivalence relation or require repeated training and testing before the relation is demonstrated, while other subjects demonstrate equivalence almost immediately. With respect to Subject 2 in this experiment, several explanations are possible. First, the subject may have had difficulty because the stimuli were purely visual. Sidman et al. (1986) suggested that delays in emergence of equivalence are more likely when purely visual stimuli are used than when stimuli are both auditory and visual. However, differences between the formation of cross-modal and one modal equivalences have not as yet been explored.

Second, Subject 2's unsolicited verbal report described the equivalence tests as containing "new words". In contrast, at the end of the first B-C post test, Subject 1 described the B-C stimuli as being "the same" as the other words; i.e., related to the A stimuli. It is possible that failure to describe the stimuli as being related contributed to failure to form the relationship.

At present the role of verbal mediation, or naming, is not fully understood. Previous research has led to the suggestion that verbal mediation may be a by-product of the equivalence relation or a function of training cross-modal relations. It is not clear whether verbal mediation is necessary for equivalence relations to emerge or whether

verbal rules sometimes emerge after equivalence relations are formed.

Third, the sequence of testing may have contributed to failure to form an equivalence relation. This interpretation is consistent with research findings presented in Chapter II. Testing in this experiment was sequenced to take advantage of the possible hierarchical nature of the relations as suggested by Stoddard and McIlvane (1986). That is, testing proceeded in the sequence of trained relations, followed by symmetrical relations, and then equivalence relations. Haywood et al. (1990) presented a different sequence that provided for rapid formation of the equivalence and reduced intersubject variability. The "simple to complex" sequence provided for testing of a symmetrical relation immediately following criterion responding on a trained relation. Equivalence tests were not presented until all other relations had been demonstrated.

Fourth, it is possible that differences in feedback between training and testing phases affected the emergence of the equivalence relation. Subject 2 frequently complained when feedback indicating correct responses was not provided, as happened during each testing phase. Instituting a social consequence produced continued responding to the computer program but only alleviated some

of his complaining. Other experimenters (e.g., Sidman & Tailby, 1982) provide for fading feedback during training so that subjects are unable to discriminate testing and training trials on the basis of feedback differences. This subject may have performed better had this methodological change been implemented.

Fifth, the subject never met the overall criterion for correct responding on the test of the symmetrical relation C-A, words matched to the parts of speech. Formation of the symmetrical relation appears to be a necessary condition for the emergence of equivalence. However, the subject did display symmetrical responding on seven of the ten word pairs tested in the C-A relation. If each set of stimuli, consisting of a word from each of class A, B and C, is functionally independent then equivalence responding might be expected to emerge for those seven stimuli.

At least two possibilities exist that might explain why symmetrical responding did not lead to equivalence relations for these seven words. First, the sets of stimuli were not independent but were interrelated through some as yet unknown feature of the experiment. This interrelatedness, if it is a feature of this experiment, does not appear to be a necessary outcome of stimulus equivalence.

Data reported by Sidman and Cresson (1973) lead to the conclusion that the sets of stimuli in their experiment were independent. The authors trained a subset of the stimuli (9 words) and then tested all of the stimuli (20 words). The testing demonstrated that the smaller set of nine equivalence relations emerged while subjects continued to respond at chance levels on the second set of words.

The second possibility is that the three unknown relations were enough to disrupt responding on the equivalence test. In either case, it would be necessary to increase correct responding on the symmetrical relation in order for equivalence relations to emerge.

Maintenance

Both subjects maintained accurate responding on the trained relation (A-B) and the symmetrical relation (B-A). For Subject 1, there was some deterioration of correct responding on tests of all four relations that included the part of speech as one member of the word pair. The subject maintained three of these relations well above baseline levels of responding. Subject 2 continued to exhibit baseline level accuracy of responding on the tests of the equivalence relations. Scores on tests of the two remaining relations deteriorated somewhat from post assessment as they had for Subject 1.

These data raise several questions that remain unanswered. For example, what variables affect the retention of trained and emergent relations? Saunders et al. (1988) speculated that the stability of a class over time may be a function of class size. That is, in a large class, each stimulus is related to more than one other stimulus in the class. If one relation is weakened, then other relations still maintain the class membership. In smaller classes this is not the case and the entire class may deteriorate. Class size may account for decreased accuracy during the maintenance phase of this experiment. However, it seems just as reasonable to speculate that maintenance is somehow related to the complexity of the stimuli. That, for example, complex subject matter such as English grammar is more difficult to maintain than simpler stimuli such as single words.

A second question that might be posed is, are relations differentially maintained? Do emergent relations maintain better than relations that have been taught? If so, under what conditions? One notable finding for Subject 1 is that one of the trained relations, (A-C) matching a part of speech to a word, deteriorated significantly in the maintenance phase. Four weeks following post assessment this subject was responding at baseline levels on the A-C test. Joyce and Wolking (1989) also tested subjects two

weeks after post assessment. They reported that one subject exhibited a slight deterioration in correct responding on a task that had been trained. Equivalence scores for this subject were higher than the score recorded on the trained relation.

The sequence of test presentations may also have affected the scores recorded during maintenance. Perhaps the A-C test set the occasion for correct responding on subsequent tests. This would be a reasonable speculation given that test sequence can enhance the emergence of equivalence relations during the acquisition phase.

Information Related to Computer Application

The computer program used in this study has the potential to be a valuable classroom tool. Subjects' performance improved on all trained relations with approximately five minutes of instruction per day. Teacher (or experimenter) intervention was unnecessary during this time. In addition, students were eager to use the computer, often requesting additional instructional time. Other students in the classroom also asked permission to work on the instructional program. In spite of the fact that they were learning something, all of the students described computer instruction as "play" and all other instruction as work.

Implications

This study was designed to evaluate the formation of equivalence classes using cognitively complex academic tasks. The results were mixed. In Experiment 1, no definitive conclusion about equivalence classes was drawn. In Experiment 2, one subject formed an equivalence class consisting of a word, a synonym, and the part of speech describing them both. The second subject failed to form an equivalence class. Successful demonstrations of stimulus equivalence in other studies add some weight to the positive results reported here (Joyce et al., 1990; Mackay, 1986; Stoddard et al., 1989). However, interpretations and recommendations from this study should be made cautiously.

The matching-to-sample procedure proved to be an effective training format. Using purely visual stimuli, all subjects accurately learned difficult symbolic relationships in three to fifteen days. The amount of training time was minimal. In general, only one training session was scheduled daily and each session lasted approximately five minutes. Training sessions were presented by computer, requiring no intervention or assistance from the teacher or experimenter. The procedure also resulted in the emergence of a number of untrained relations (demonstration of stimulus equivalence) for one subject in Experiment 2. Much of the trained and emergent behavior maintained two and four

weeks following the post assessment with no review of material.

The computer application proved to be highly reinforcing. Experimental subjects and other students in the classroom frequently requested instructional time and complained when their time was completed. Further indications of the highly motivating nature of the program were the changing classroom and home behaviors of the subjects. The subjects who participated in this experiment had been identified by the teacher as unmotivated to learn the English language. However, informal interviews with the teacher and parents after the experiment indicated that the subjects often discussed what they had learned on the computer and sometimes described training stimuli and asked for the correct answers.

Based upon the preceding discussion, the matching-to-sample procedure, arranged so that equivalence classes may be formed, can be recommended as an efficient and productive method of teaching cognitively complex academic tasks. However, the application of the stimulus equivalence paradigm can be improved. First, the probability that an equivalence relation will emerge may be improved by sequencing testing and training so that symmetrical relations are verified before equivalence relations are tested. In addition, feedback during training sessions

should be faded in order to avoid discriminations between testing and training procedures. Once a class of stimuli has been established, review sessions appear to be necessary in order to maintain accurate responding.

Future Research

While it is clear that stimulus equivalence can lead to improved teaching practices, many procedural and process issues must be resolved before it can be optimally applied. In order to continue to address complex academic subject matter, a number of issues must be resolved that relate to the structure of the matching-to-sample procedure. For example, the number and variability of the distractor stimuli became a problem during this study. This problem generated several questions that were not answered by data but by custom. Questions that should be asked include: How many distractor stimuli are necessary? Is the number different during testing and training? Does learning actually occur during testing if the distractor stimuli are varied? If the distractor stimuli are drawn from a different set that are never taught, does learning occur?

A second issue relates to the testing and training sequence. Test sequence has been identified as an important variable during post assessment that is just beginning to be studied (Haywood et al., 1990; Stoddard & McIlvane, 1986). However, test sequence may also be important during

pretests. Sequencing tests from simple to complex seems to enhance the formation of equivalence classes. However, other test sequences have also resulted in the emergence of symmetrical and equivalence relations. It has been suggested that the test itself sets the occasion for a relation to emerge. If this is the case, testing may enhance or block learning during pretesting as well as during postassessment.

A third issue concerns the maintenance of relations following postassessment. In the present study, relations deteriorated differentially after two and four weeks. That is, some relations maintained and slightly improved, some relations deteriorated after two weeks and then improved at the four week follow-up, and one trained relation deteriorated to baseline levels. More and longer term studies of maintenance need to be completed. In addition, studies are needed that address the suggestion that relations in a larger class may maintain better than those in a smaller class.

Finally, research should continue to address stimulus equivalence using cognitively complex academic subject matter as stimuli. In the current study, both stimuli (Spanish/English words and English grammar) and subjects were complex and varied from previous research. Subjects were primarily Spanish speaking, unmotivated, and diagnosed

with learning disabilities. Results may be related to the stimuli, the subjects, or an interaction between the two. Future research can address cognitively complex academic subject matter using a less complex population. History, science, math, language, geography and music are examples of fields that require the formation of classes of stimuli that are equivalent. Questions may address whether and what equivalence classes can be formed, stimulus class expansion and contextual control of stimuli.

A more complete understanding of the development of equivalent stimulus classes may move us closer to an effective technology of instruction. The matching-to-sample paradigm, desk top computer, and hyper media programming tools may make the implementation of such a technology both effective and efficient.

GLOSSARY

Cognitively Complex Academic Subject Matter--subject matter typically presented later in the school curriculum. These include content area subject matter such as chemistry, geometry, English grammar, etc.

Cross-modal Training--training stimuli in one mode (e.g., auditory) to stimuli in different mode (e.g., visual).

Latency--the period of time that elapses between the presentation of a stimulus and the response that it controls.

Maintenance--retaining the necessary level of fluency once instruction is terminated.

Matching-to-Sample--presentation of a sample and a number of comparison stimuli in which a comparison is to be matched to the sample. Similar stimuli are assumed to be identical whereas dissimilar stimuli are assumed to be equivalent.

Reflexivity--generalized identity matching.

Stimulus Generalization--responding in a similar manner to new stimuli that have similar physical properties as a previous stimulus or class of stimuli

Symmetry--reversibility of sample and comparison in a matching-to-sample task.

Stimulus Equivalence--a stimulus-stimulus relationship defined by the presence of the three properties: reflexivity, symmetry, and transitivity.

Transitivity--the appropriate recombination of samples and comparisons into a new conditional discrimination.

REFERENCES

Behrmann, M. M., & Lahm, L. (Eds.). (1984). Proceedings of the national conference on the use of microcomputers in special education, Reston, VA: Council for Exceptional Children.

Behrmann, M. M., Stevens, S., & Levy, S. A. (1984). Future of computers in special education. In Behrmann, M. (Ed.), Handbook of microcomputers in special education, (pp. 243-269). San Diego, CA: College-Hill Press.

Bernal, E. (1983). Trends in bilingual special education. Learning Disability Quarterly, 6, 424-431.

Billings, K. (1983). Developing mathematical concepts with microcomputer activities. Arithmetic Teacher, 30, 18-19,57-59.

Bush, K., Sidman, M., & deRose, T. (1989) Contextual control of emergent equivalence relations. Journal of the Experimental Analysis of Behavior, 51, 29-46.

Cumming, W. W., & Berryman, R. (1965). The complex discriminated operant: Studies on matching-to-sample and related problems. In D. I. Mostofsky (Ed.), Stimulus generalization (pp. 284-330). Stanford, CA: Stanford University Press.

Cummins, A. (1983). Bilingualism and special education: Program and pedagogical issues. Learning Disability Quarterly, 6, 373-386.

Edwards, J. (1974). Direct measurement of the ways children know language and numbers. In J. L. Sheridan (Ed.), Readings for experimental psychology, New York: Holt, Rinehart, and Winston.

Englert, C. S. (1984). Effective direct instruction practices in special education settings. Remedial and Special Education, 5, 38-47.

Garner, J. B. & Campbell, P. H. (1987). Technology for persons with severe disabilities: Practical and ethical considerations. The Journal of Special Education, 21, 122-132.

Gast, D. L., VanBiervliet, A., & Spradlin, J. E. (1979). Teaching number-word equivalences: A study of transfer. American Journal of Mental Deficiency, 83, 524-527.

Glatthorn, A. A. & Rosen, B. C. (1988). McDougal, Littell English - Green Level, Evanston, IL: McDougal, Littell.

Goldstein, M. (1985). Matrix and stimulus equivalence training. In S. F. Warren & A. K. Rogers-Warren (Eds.), Teaching functional language - Generalization and maintenance of language skills, Baltimore, MD: University Park Press.

Hannaford, A. E. (1987). Computers and exceptional individuals. In J. D. Lindsey (Ed.), Computers and exceptional individuals (pp. 1-22). Columbus, OH: Merrill Publishing.

Haring, T. G., Breen, C. G., & Laitinen, R. E. (1989). Stimulus class formation and concept learning: Establishment of within- and between-set generalization and transitive relationships via conditional discrimination procedures. Journal of the Experimental Analysis of Behavior, 52(1), 13-26.

Haywood, C. S., Fields, L., Adams, B. J., & Verhave, T. (May, 1990). Test order and the regularity of equivalence class formation. Paper presented at the Sixteenth Annual Conference of the Association for Behavior Analysis: International, Nashville, TN.

Hofmeister, A. M. (1982). Microcomputers in perspective. Exceptional Children, 49, 115-121.

Joyce, B. G, Joyce, J. H., & Wellington, B. (May, 1990). Stimulus equivalence: An approach for training a foreign language to children with traumatic brain injury. Paper presented at the Sixteenth Annual Conference of the Association for Behavior Analysis: International, Nashville, TN.

Joyce, B. G. & Wolking, W. D. (1989). Stimulus equivalence: An approach for teaching beginning reading skills to young children. Education and Treatment of Children, 12(2), 109-122.

Lazar, R. M., Davis-Lang, D., & Sanchez, L. (1984). The formation of stimulus equivalences in children. Journal of the Experimental Analysis of Behavior, 41, 251-266.

Mackay, H.A. (1986). Stimulus equivalence in rudimentary reading and spelling. Analysis and Intervention in Developmental Disabilities, 5, 373-387.

Mackay, H. A. & Sidman, M. (1984). Teaching new behavior via equivalence relations. In P. H. Brooks, R. Sperber, & McCauley, C. (Eds.), Learning and cognition in the mentally retarded. (pp. 493-513). Hillsdale, NJ: Lawrence Erlbaum Assoc.

McCarthy, M. (1980). Minimum competency testing and handicapped students. Exceptional Children, 47, 166-173.

McDonagh, E.C., McIlvane, W. J., & Stoddard, L. T. (1984). Teaching coin equivalences via matching to sample. Applied Research in Mental Retardation, 5, 177-197.

Mercer, C. D. & Mercer, A. R. (1981). Teaching students with learning problems. Columbus OH: C. E. Merrill Publishing.

Reeve, K., Fields, L., Adams, B. J., Kaplan, H., & Goss, M. (May, 1990). The generalization of equivalence relations. Paper presented at the Sixteenth Annual Conference of the Association for Behavior Analysis: International, Nashville, TN.

Rogerson, H. D., Esarey, G., Jasnow, C., Hershelman, S. T., Schmandt, L. M., Smith, D. A., & Snellings, C. M. (1988). Words for students of English - Volume 6, Pittsburgh, PA: University of Pittsburgh Press

Saunders, R. R., Saunders, K. J., Kirby, K. C., & Spradlin, J. E. (1988). The merger and development of equivalence classes by unreinforced conditional selection of comparison stimuli. Journal of the Experimental Analysis of Behavior, 50, 145-162.

Sidman, M. (1971). Reading and auditory-visual equivalences. Journal of Speech and Hearing Research, 14, 5-13.

Sidman, M. (1988) Functional analysis of emergent verbal classes. In T. Thompson & M. Zeiler, (Eds.). Units of analysis and integration of behavior. (pp. 213-245). Hillsdale, NJ: Lawrence Erlbaum Associates.

Sidman, M. & Cresson Jr., O. (1973). Reading and crossmodal transfer of stimulus equivalences in severe retardation. American Journal of Mental Deficiency, 77, 515-523.

Sidman, M., Cresson, Jr., O. & Willson-Morris, M. (1974). Acquisition of matching to sample via mediated transfer. Journal of the Experimental Analysis of Behavior, 22, 261-273.

Sidman, M., Kirk, B. & Willson-Morris, M. (1985). Six-member stimulus classes generated by conditional-discrimination procedures. Journal of the Experimental Analysis of Behavior, 43, 21-42.

Sidman, M., Rauzin, R. Lazar, R., Cunningham, S., Tailby, W. & Carrigan, P. (1982). A search for symmetry in the conditional discriminations of rhesus monkeys, baboons, and children. Journal of the Experimental Analysis of Behavior, 37, 23-44.

Sidman, M., & Stoddard, L. T. (1967). The effectiveness of fading in programming simultaneous form discrimination for retarded children. Journal of the Experimental Analysis of Behavior, 10, 3-15.

Sidman, M. & Tailby, W. (1982). Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. Journal of the Experimental Analysis of Behavior, 37, 5-22.

Sidman, M., Wilson-Morris, M. & Kirk, B. (1986). Matching-to-sample procedures and the development of equivalence relations: The role of naming. Analysis & Intervention in Developmental Disabilities, 6, 1-19.

Spradlin, J. E., Cotter, V. W., & Baxley, N. (1973). Establishing a conditional discrimination without direct training: A study of transfer with retarded adults. American Journal of Mental Deficiency, 77, 556-566.

Spradlin, J.E. & Saunders, R.R. (1984). Behaving appropriately in new situations: A stimulus class analysis. American Journal of Mental Deficiency, 88, 574-579.

Spradlin J.E. & Saunders, R.R. (1986). The development of stimulus classes using match-to-sample procedures: Sample classification versus comparison classification. Analysis & Intervention in Developmental Disabilities, 6, 41-58.

Stoddard, L.T., Brown, J., Hurlbert, C., Manoli, C., & McIlvane, W. J. (1989). Teaching money skills through stimulus class formation, exclusion, and component matching methods: Three case studies. Research in Developmental Disabilities, 10, 413-439.

Stoddard, L. T. & McIlvane, W. J. (1986) Stimulus control research and developmentally disabled individuals. Analysis and Intervention in Developmental Disabilities, 6, 155-178.

Tawney, J. W. & Sniezek, M. (1985). Education programs for severely mentally retarded elementary-age children: Progress, problems and suggestions. In D. Bricker & J. Filler (Eds.), Severe Mental Retardation, (pp 76-96). Reston, VA: Division of Mental Retardation of the Council for Exceptional Children.

VanBiervliet, A. (1977). Establishing words and objects as functionally equivalent through manual sign training. American Journal of Mental Deficiency, 82, 178-186.

Wang, M. (1987) Toward achieving educational excellence for all students: Program design and student outcomes. Remedial and Special Education, 8(3), 25-34.

Wiebe, J. H. (1987). The software domain. In J. D. Lindsey (Ed.), Computers and exceptional individuals (pp. 37-59). Columbus, OH: Merrill Publishing Co.

Wulz & Hollis (1979). Manual signing and the development of reading skills. In L. Schiefelbusch and J. H. Hollis (Eds.), Language Intervention from Ape to Child (pp. 467-488). Baltimore, MD: University Park Press.

APPENDIX A
PERMISSION AND INFORMED CONSENT FORMS

APPLICATION FOR RESEARCH IN ALACHUA COUNTY SCHOOLS

Applicant: Jolenea Stoutimore Phone: 392-0755 Date: 9/14/89

Address: G315 Norman Hall, U of Fl., Gainesville, Fl. 32611

Applicant is: Doctoral Student

Title of Research Proposal: Computer Assisted Second Language Instruction

Purpose of Research: This is a demonstration project to show that non-English speaking students can efficiently acquire English vocabulary and related skills using computer assisted instruction

Brief Summary of Proposal: This study will present words that are to be matched with pictures, sounds and other words or phrases in a microcomputer program that is game-like in format.

Population Needs: 3-4 Subjects Students whose first language is Spanish and who have a limited English vocabulary.

Data Needed: I.Q. test scores, if available, and any language competency test results (both English and Spanish.

Total Time Required: 10 to 15 sessions, 20 to 30 minutes in length.

School Requested: Lincoln Middle School.

If this application is approved, I agree to observe all legal requirements regarding the use of research and submit a final copy of the research report to the Alachua County School Board.

UNIVERSITY OF FLORIDA
Institutional Review Board

August 28, 1989

TO: Dr. William D. Wolking
G315 NRN

FROM: C. Michael Levy, Chair,
University of Florida Institutional Review
Board

SUBJECT: Approval of Project # 89.206
Stimulus Equivalence and second language
learning

I am pleased to advise you that the University of Florida Institutional Review Board has recommended the approval of this project. The Board concluded that your subjects will not be placed at risk in this research, and it is essential that you obtain legally effective informed consent from each participant's parent or legal guardian. When it is feasible, you should obtain signatures from both parents.

If you wish to make any changes in this protocol, you must disclose your plans before you implement them so that the Board can assess their impact on your project. In addition, you must report to the Board any unexpected complications arising from the project which affect your subjects.

If you have not completed this project by August 28, 1990, please telephone our office (392-0433) and we will tell you how to obtain a renewal.

By a copy of this memorandum, your chair is reminded of the importance of being fully informed about the status of a project involving human subjects in your department, and for reviewing these projects as often as necessary to insure that each project is being conducted in the manner approved by this memorandum.

PARENTAL INFORMED CONSENT FORM

Your chid is being asked to participate in a study designed to teach English as a second language. Each child in the study will participate in 10 to 15 sessions that are approximately 20 minutes in length. Each session will be administered by a UF doctoral student, Jolenea Stoutimore. All sessions will be scheduled during individual study times.

The study will present words that are to be matched with pictures, sounds and other words or phrases. The words, pictures, and sounds will be presented in a microcomputer program that is game-like in format. Pictures and/or words are presented on the computer screen. Matches are then selected by moving a pointer onto the chosen item and pressing the smith on a Macintosh mouse.

Students participating in this study should learn vocabulary and word meanings in a second language and gain experience with computer-assisted instruction while engaging in an enjoyable learning experience. Although this procedure has not previously been used to teach a second language, other studies have demonstrated that it is an effective way to teach a variety of academic skills, including reading comprehension and vocabulary.

The information obtained will be kept confidential. Your child's name will not be used in the written results. There is no risk involved and there will be no monetary compensation. You, or your child, have the right to withdraw consent at any time without prejudice.

Participation in this study is completely voluntary and is not a class requirement. Participation or non-participation will not affect your child's grade. Students who do participate will not receive extra credit or bonus points, although they should benefit from the learning experience.

If you have any questions about this study, please contact me at work or home. I will gladly discuss this information in more detail.

I have read and I understand the procedure above. I agree to allow my child ________________ to participate in the procedure and I have a copy of this description.

Parent/Guardian Signature   Date

Script for Verbal Permission from Student

"Hello, (student name), I'm Jolenea. I have a game on the computer that will help you learn some new words.
Would you like to play?"

APPENDIX B
ENGLISH/SPANISH WORD LISTS

WORD LISTS

English	Spanish	English	Spanish
airplane	avion	corn	maiz
apple	manzana	custodian	conserje
arm	brazo	deer	venado
arrow	fleche	dessert	postre
axe	hacha	doll	muneca
baby	nino/nina	door	puerta
bacon	tocino	dress	vestido
bakery	panderia	drum	tambor
ball	bola	duck	pato
balloon	globo	egg	heuvo
barber	peluquero	engine	maquina
barn	troje	face	cara
basket	cesto	factory	fabrica
beach	playa	fall	otono
bed	cama	farmer	ranchero
bedroom	dormitoria	father	padre
blackboard	pizarra	faucet	llave
boat	barco	fence	cerca
bonfire	fogon	fender	guardabarros
boy	muchacho	fish	pez
branch	rama	floor	piso
bread	pan	food	comida
brick	ladrillo	forest	bosque
bridge	puente	fox	zorro
brother	hermano	friend	amigo
bubble	burbuja	furniture	muebles
bush	arbusto	garbage	basura
butcher	carnicero	garden	jardin
cab	taxi	gate	porton
can (tin)	lata	girl	muchacha
candy	dulce	glass	vaso
canoe	chalupa	glove	guante
car	coche	goat	cabra
cashier	cajero	grandmother	abuela
cat	gato	grandfather	abuelo
cheese	queso	grass	zacate
chef	cocinero	hair	pelo
chicken	pollo	hammer	martillo
child	hijo	hat	sombrero
church	iglesia	head	cabeza
clock	reloj	heart	corozon
clothing	ropa	helmet	yelmo
cloud	nube	hen	gallina
coat	saco	honey	miel
cup	taza	horn	bocina

WORD LIST (CONT.)

English	Spanish	English	Spanish
horse	caballo	pool	piscina
hose	manguera	postman	cartero
house	casa	potato	papa
jacket	chaqueta	puppet	titere
kitchen	cocina	rack	toallero
kite	papalote	raccoon	mapache
kitten	gatito	rain	lluvia
knife	cuchillo	rake	rastro
ladder	escalera	river	rio
lake	lago	rocket	cohete
leaf	hoja	root	raiz
letter	carta	rug	tapete
lettuce	lechuga	ruler	regla
librarian	bibliotecario	sausage	salchicha
library	biblioteca	saw	sierra
lumber	madera	school	escula
lock	cerradura	scooter	monopatin
magazine	revista	seed	semilla
market	mercado	seamstress	costurera
match	fosforo	shelf	estante
meat	carne	shovel	pala
milk	leche	sink	lavado
mirror	espejo	sister	hermana
money	dinero	skate	patin
moon	luna	skillet	sarten
mouse	cajita	skirt	falda
mouth	boca	spring	primavera
movies	pelicula	spoon	cuchara
mushroom	seta	star	estrella
mustard	mostaza	stem	tallo
nail	clavo	store	tienda
needle	aguja	storekeeper	tendero
newspaper	periodico	street	calle
notebook	cuarderno	summer	verano
nurse	enfermera	sun	sol
orange	naranja	swing	colummpio
owl	lechuza	table	mesa
paint	pintura	tablecloth	mantel
paper	papel	tree	arbol
party	fiesta	truck	camion
pig	cochino	utensils	cubierto
pin	alfiler	vase	florero
pineapple	pina	waiter	camarero
pliers	aliates	wallet	cartera
plumber	plomero	water	agua

WORD LISTS (CONT.)

English	Spanish
windshield	parabrisa
winter	invierno
window	ventana
wiper	limpia
witch	bruja
world	mundo
worker	trabajador

APPENDIX C
MOUSE PRETRAINING

1. Instructions to subjects:

   The box on the table is called a mouse. See what happens when you move it.

2. Once the subject is moving the mouse around the screen, new instructions are given:

   You see a window in the center of the screen. Move the hand so that it is in the window.

Repeat step 2 until the subject has made 5 correct responses.

3. Subjects are instructed to move the cursor into the center window:

   Move the hand into the center window. Now push on this button (experimenter points to the mouse button).

4. When the center window is selected, a new screen appears on which there are four windows displayed in the same configuration as that used during the experiment. Subjects are instructed to move the hand into one of the boxes and push on the button. Steps 3 and 4 are repeated until the subject makes 5 correct responses on both steps.

   A correct response consists of correct cursor placement an pushing the button when the cursor is inside one of the boxes.

5. Verbal praise, e.g., "good", is given for each correct response.

APPENDIX D
EXPERIMENT 2: SEQUENCE OF TESTING AND TRAINING WITHIN EACH RELATION

Subject 1:

Baseline 1 & 2	Training	Post Assessment	Maintenance 1 & 2
B-C	A-B	A-B	A-B
C-B	A-C	B-A	B-A
A-B		A-C	A-C
B-A		Train A-C	C-A
A-C		C-A	B-C
C-A		B-C	C-B
		C-B	
		B-C	

Subject 2:

Baseline 1 & 2	Training	Post Assessment	Maintenance 1 & 2
B-C	A-B	A-B	A-B
C-B	A-C	B-A	B-A
A-B		A-C	A-C
B-A		C-A	C-A
A-C		B-C	B-C
C-A		C-B	C-B
		Train A-C	
		A-B	
		A-C	
		B-C	
		C-B	
		Train A-B	
		Train A-C	
		B-C/A-C mixed	

BIOGRAPHICAL SKETCH

Jolenea Ferro was born in Kansas City, Missouri, where she attended elementary and high schools. She graduated from the University of Missouri at Kansas City in 1968, receiving a B.A. in psychology.

Since 1968, her education and experience have been concentrated in the area of behavior analysis. Her master's program involved the use of experimental behavior analysis in both laboratory and applied settings. These included fetal and infant development, language development, and social and academic behaviors. She earned her M. A. in Psychology from the University of Missouri at Kansas City in 1972.

In 1988, Jolenea began full time doctoral study in special education at the University of Florida. Her program has been concentrated on the instructional use of microcomputers, applied behavior analysis, and teacher education. Upon completion of the Ph.D. program, she will continue to teach, conduct research, and write.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Wm. D. Wolking

William D. Wolking, Chairman
Professor of Special Education

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Cary L. Reichard

Cary L. Reichard
Professor of Special Education

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

William R. Reid

William R. Reid
Professor of Special Education

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Elroy J. Bolduc

Elroy J. Bolduc
Professor of Instruction and
Curriculum

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Brian Iwata
Professor of Psychology

This dissertation was submitted to the Graduate Faculty of the College of Education and to the Graduate School, and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August, 1990

Dean, College of Education

Dean, Graduate School